

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 17, 2003, 08:44:21 ; Search time 450.327 Seconds  
(without alignments)  
15572.506 Million cell updates/sec

Title: US-09-766-511b-51

Perfect score: 3114  
Sequence: 1 cttaatgttggaagtctctt.....tttaaaaaaaaaaaaaaa 3114

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002.\*

	1:	2:	3:	4:	5:	6:	7:	8:	9:	10:	11:	12:	13:	14:	15:	16:	17:	18:	19:	20:	21:	22:	23:	24:
	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1980.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1981.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1982.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1983.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1984.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1985.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1986.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1987.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1988.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1989.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1990.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1991.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1992.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1993.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1994.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1995.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1996.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1997.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1998.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA1999.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA2000.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA2001A.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA2001B.DAT.*	/SID52/gcgcdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3114	100.0	3114	22	AAS01375 Human TANGO 405 cD
2	591.4	19.0	850	20	AAS01375 Human SDCMP3 poly
3	379.2	12.2	1227	19	AAV42549 Mouse dectin-2 cDN
4	379.2	12.2	1252	22	AAS01378 Murine TANGO 405 a
5	376.2	12.1	630	20	AAZ07532 Mouse SDCMP3 poly
6	364.6	11.7	3708	20	AAZ77530 Human ovarian tumo
7	364.6	11.7	3748	21	AAC77686 Human cancer assoc
8	356.2	11.4	3647	24	ABK83658 Human cDNA differe
9	351.4	11.3	803	22	AAI97506 Human neuroblastom

10	345.8	11.1	821	22	AAS01376 Murine TANGO 405 c
11	317.6	10.2	56583	21	AAF21125 Human low adenosin
12	317.6	10.2	56583	21	AA335003 Human adenosine re
13	303.2	9.7	501	19	AAV42554 Mouse dectin-2 ext
14	283.8	9.1	1312	22	AAF90241 Nucleotide sequenc
15	272.8	8.8	827	22	AAD19729 Dendritic cell (DC
16	269	8.6	800	22	AAD19730 Dendritic cell (DC
17	257.8	8.3	3298	17	AA45082 Beta-1-4-galactosy
18	255.4	8.2	684	22	AAF98678 Human ovarian canc
19	254.2	8.2	324	22	AAK63162 Human immune/haema
20	244.4	7.2	3929	24	AB190324 Mouse ischaemic co
21	201.2	6.5	444	24	ABK52914 Human cDNA encodin
22	200.4	6.4	402	24	ABK52903 Human cDNA encodin
23	197	6.3	1091	22	AAS31385 Human polynucleoti
24	197	6.3	1091	24	ABQ66709 Human cDNA encodin
25	197	6.3	1096	22	AAS31224 Human polynucleoti
26	197	6.3	1096	24	ABQ66548 Human polynucleoti
27	197	6.3	1104	20	AA04865 Primate DCMPI C-le
28	197	6.3	1307	24	AB190686 Human polynucleoti
29	179.4	5.8	582	24	ABK52904 Human dendritic ce
30	177.8	5.7	753	24	ABK52906 Human dendritic ce
31	169.6	5.4	1013	24	ABK52905 Human dendritic ce
32	154.8	5.0	558	24	ABK52907 Human dendritic ce
33	144.8	4.6	561	24	ABK52911 Human dendritic ce
34	140	4.5	448	24	ABK52912 Human dendritic ce
35	140	4.5	666	24	ABK52910 Human dendritic ce
36	137.4	4.4	568	24	ABK52908 Human dendritic ce
37	137.4	4.4	1418	20	AA04867 rodent DCMPI C-lec
38	130.2	4.2	403	24	ABK52909 Human dendritic ce
39	118.6	3.8	375	16	AAT24623 Human gene signatu
40	117.4	3.8	968	22	AAF81743 Human membrane ass
41	117.4	3.8	997	20	AA52974 Protein PRO244 cDN
42	117.4	3.8	997	22	AA545936 Human DNA encoding
43	117.4	3.8	997	22	AAF72432 Human PRO244 cDNA.
44	111.8	3.6	458	22	AA524836 Human ovarian PCR-
45	111.8	3.6	567	22	AAH83473 Human ovarian tumo

#### ALIGNMENTS

#### RESULT 1

AAS01375  
ID AAS01375 standard; cDNA; 3114 BP.

XX AAS01375;

XX AC AAS01375;

XX DT 04-JUL-2001 (first entry)

XX DE Human TANGO 405 cDNA sequence.

XX DE Human TANGO 210; clone jthLal52h06; TANGO 364; TANGO 366; dectin-2;

KW INTERCEPT 394; INTERCEPT 400; TANGO 405; cellular process regulator;

KW gene therapy; growth modulator; proliferation; cell differentiation;

KW lymphocyte; bone marrow cell migration; leukaemia; lymphoma;

KW autoimmune disorder; ss.

XX OS Homo sapiens.

XX PH Key

XX CDS Location/Qualifiers

FT 154..783

FT /\*tag= a

FT /product= "TANGO 405 protein"

FT /note= "The ORF is specifically claimed"

FT 154..297

FT /\*tag= b

FT 298..780

FT /\*tag= c

XX WO200118016-A1.

XX PD 15-MAR-2001.

XX

PF 30-JUN-2000; 2000WO-US18174.  
 XX  
 PR 10-SEP-1999; 99US-0393996.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Fraser CC, Sharp JD, Wrighton N, Myers PS, Goodearl ADJ;  
 XX  
 DR WPI; 2001-183280/18.  
 XX P-PSDB; AAU00479.  
 XX  
 PT Isolated nucleic acid molecules encoding proteins useful as modulating  
 PT agents in regulating a variety of cellular processes are used for  
 PT treating e.g. cancer and autoimmune disorders -  
 XX  
 PS Claim 2; Fig 6A-6C; 326pp; English.  
 XX  
 CC The present sequence encoding for human TANGO 405 protein is isolated  
 CC from cDNA clone jthla152h06 from a human mixed lymphocyte reaction cDNA  
 CC library. It is 1 of 6 novel human proteins which include TANGO 210  
 CC (AAU00469), TANGO 364 (AAU00471), TANGO 366 (AAU00472), INTERCEPT 394  
 CC (AAU00473) and INTERCEPT 400 (AAU00476). Novel sequences for murine  
 CC TANGO 210 (AAU00470), INTERCEPT 400 (AAU00477), TANGO 405 (AAU00480) and  
 CC a rat INTERCEPT 400 (AAU00478) sequence are also described. The nucleic  
 CC acids encoding these novel proteins are useful as modulating agents in  
 CC regulating a variety of cellular processes and can be used to express  
 CC the proteins in a host cell in gene therapy applications. Human and  
 CC murine TANGO 405 proteins show sequence homology to murine dectin-2.  
 CC TANGO 405 modulates growth, proliferation, survival, differentiation,  
 CC activity, morphology and movement/migration of human lymphocytes and  
 CC bone marrow cells and tissues and can be used to prevent, diagnose or  
 CC treat leukaemia, lymphomas and autoimmune disorders.  
 XX  
 XX Sequence 3114 BP; 1001 A; 527 C; 517 G; 1069 T; 0 other;  
 XX

Query Match 100.0%; Score 3114; DB 22; Length 3114;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 3114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTAATGTTGGAAGTCCTTAGTCCTATGAGAGTGTAGCAGTTGTGCCCTGAGCTCTA 60  
 DB 1 CTTAATGTTGGAAGTCCTTAGTCCTATGAGAGTGTAGCAGTTGTGCCCTGAGCTCTA 60

QY 61 GCTTCTTTAAATGAAGTCAGTCTCTGGCAACATCTTTAGGAGAGAGGTACAAAGGT 120  
 DB 61 GCTTCTTTAAATGAAGTCAGTCTCTGGCAACATCTTTAGGAGAGAGGTACAAAGGT 120

QY 121 TCCTGGACCTTCTCAACACAGGGAGGCTGCATATATGATGCAAGAGCAGCAACCTCAAAGT 180  
 DB 121 TCCTGGACCTTCTCAACACAGGGAGGCTGCATATATGATGCAAGAGCAGCAACCTCAAAGT 180

QY 181 ACAGAGAAAGAGCTGGTGTCTCCCTGAGACCTGCTCTGTGGTGGGATTTCCATTGCA 240  
 DB 181 ACAGAGAAAGAGCTGGTGTCTCCCTGAGACCTGCTCTGTGGTGGGATTTCCATTGCA 240

QY 241 CTCCTCAGTGTCTCTCATTTGAGCTGTGTAGTAACCTTACCATTACATATGGTGCA 300  
 DB 241 CTCCTCAGTGTCTCTCATTTGAGCTGTGTAGTAACCTTACCATTACATATGGTGCA 300

QY 301 ACTGGCAAAAGGCTGTCTGAACCTACACTCATATCATTCATCAAGTCCTCACTGCTCAGTGA 360  
 DB 301 ACTGGCAAAAGGCTGTCTGAACCTACACTCATATCATTCATTCATCAAGTCCTCACTGCTCAGTGA 360

QY 361 GGGCAAAAGGTCGAGGCTGGGATGTGGCCAGCTCTTGGAAAGTCATTGGTTCCAGT 420  
 DB 361 GGGCAAAAGGTCGAGGCTGGGATGTGGCCAGCTCTTGGAAAGTCATTGGTTCCAGT 420

QY 421 TGCTACTTCATTTCCTGAGAGAGAGGTTTGGTCTTAAGAGTGAGCAAGACTGTGTGAG 480  
 DB 421 TGCTACTTCATTTCCTGAGAGAGAGGTTTGGTCTTAAGAGTGAGCAAGACTGTGTGAG 480

QY 481 ATGGAGCACATTGGTTGTTCACACAGAGAGCAGCAATTCATTGTCCAGAG 540  
 DB 481 ATGGAGCACATTGGTTGTTCACACAGAGAGCAGCAATTCATTGTCCAGAG 540

DB 481 ATGGAGCACATTGGTTGTTCACACAGAGAGCAGCAATTCATTGTCCAGAG 540  
 QY 541 CTGAATGAGTCATTTTCTTATTTCTGGGGCTTTTCAGACCCACAAGGTAAATAATTTGG 600  
 DB 541 CTGAATGAGTCATTTTCTTATTTCTGGGGCTTTTCAGACCCACAAGGTAAATAATTTGG 600  
 QY 601 CAATGATGTATGATAGACACCTTATGAGAAAATGTCAGATTTTGGCACCTAGGTGAGCC 660  
 DB 601 CAATGATGTATGATAGACACCTTATGAGAAAATGTCAGATTTTGGCACCTAGGTGAGCC 660  
 QY 661 AATCATCTCGAGACCAATGCTTCAATAGTCTTCTGGAAACCTACAGGATGGGGCTGG 720  
 DB 661 AATCATCTCGAGACCAATGCTTCAATAGTCTTCTGGAAACCTACAGGATGGGGCTGG 720  
 QY 721 AATGATGTATCTGTGAAACTAGAGGAATTCATATGTGAGATGAATTAAGATTTACCTA 780  
 DB 721 AATGATGTATCTGTGAAACTAGAGGAATTCATATGTGAGATGAATTAAGATTTACCTA 780  
 QY 781 TGAGTAGAGCTTAATTCGAAAGAGAGAGAAATTAAGTGAATTTTCCCTGAGCT 840  
 DB 781 TGAGTAGAGCTTAATTCGAAAGAGAGAGAAATTAAGTGAATTTTCCCTGAGCT 840  
 QY 841 CTTTAAATTTGAACCTATCATGAATGATAATTTCTTCTGAAATTTACACATAATTCCTT 900  
 DB 841 CTTTAAATTTGAACCTATCATGAATGATAATTTCTTCTGAAATTTACACATAATTCCTT 900  
 QY 901 ATGTTATAGAGTTCACAGAAATGGAAGATACCTGTTCCTTAAATCAATCTTCTCGT 960  
 DB 901 ATGTTATAGAGTTCACAGAAATGGAAGATACCTGTTCCTTAAATCAATCTTCTCGT 960  
 QY 961 TTCCCTCTTTCCATTAATGATGAGATGCAACCTTCCCTCTCTTTGTTCCCATTTCTTCACTT 1020  
 DB 961 TTCCCTCTTTCCATTAATGATGAGATGCAACCTTCCCTCTCTTTGTTCCCATTTCTTCACTT 1020  
 QY 1021 GTATATCATTTTCTTCTTCCACTTACATACACAAATATTTATTTGTTCCAGAGCT 1080  
 DB 1021 GTATATCATTTTCTTCTTCCACTTACATACACAAATATTTATTTGTTCCAGAGCT 1080  
 QY 1081 GTACTATTTGTTGTTAGAAAGATTTAAGGCGAGTATCTTTTGAAGATTTATGACTTTCC 1140  
 DB 1081 GTACTATTTGTTGTTAGAAAGATTTAAGGCGAGTATCTTTTGAAGATTTATGACTTTCC 1140  
 QY 1141 TTCCCTCAATATACCATTAAGAAATCTTTTGGTCAAGATGGTAGTTGGAGACTACAAATCAT 1200  
 DB 1141 TTCCCTCAATATACCATTAAGAAATCTTTTGGTCAAGATGGTAGTTGGAGACTACAAATCAT 1200  
 QY 1201 CTGAAGGCTGACAGAGTGTGAAAGACATGTTTCTAGATGGCTCACTCAGATGGCTGGC 1260  
 DB 1201 CTGAAGGCTGACAGAGTGTGAAAGACATGTTTCTAGATGGCTCACTCAGATGGCTGGC 1260  
 QY 1261 AACTTGGTGTGGCTTATTAATGTAACCTGGAATAAATTTTATTTCTGCAAGTTAGGATTT 1320  
 DB 1261 AACTTGGTGTGGCTTATTAATGTAACCTGGAATAAATTTTATTTCTGCAAGTTAGGATTT 1320  
 QY 1321 GGCATTTTATATATGTTGATTCAATCAAGTTTGGCAAGAGGCTTTCGATACTGCTATA 1380  
 DB 1321 GGCATTTTATATATGTTGATTCAATCAAGTTTGGCAAGAGGCTTTCGATACTGCTATA 1380  
 QY 1381 TCCTGTATTTCTGTTTATTTGTTTATTTCTGAGAAATATGTTAAGATCTCTCGCTG 1440  
 DB 1381 TCCTGTATTTCTGTTTATTTGTTTATTTCTGAGAAATATGTTAAGATCTCTCGCTG 1440  
 QY 1441 ATTGGAAATTTGCTTATTTCTCATTTAAATTTTGTCAAACTTTCTTGTGTCGAAGCAT 1500  
 DB 1441 ATTGGAAATTTGCTTATTTCTCATTTAAATTTTGTCAAACTTTCTTGTGTCGAAGCAT 1500  
 QY 1501 TTCCTGTTTACCAATCACTTATCTGAAATATGATGGTTAGCAAGTTGAGATA 1560  
 DB 1501 TTCCTGTTTACCAATCACTTATCTGAAATATGATGGTTAGCAAGTTGAGATA 1560  
 QY 1561 ACTAGAGCTGTAAATCACTTATTTAAATGGCAATGATGACAGTTTATTTTATGTT 1620  
 DB 1561 ACTAGAGCTGTAAATCACTTATTTAAATGGCAATGATGACAGTTTATTTTATGTT 1620

QY 1621 ATATATAAACCTCAACAATTTTCCAAACAATTTACCAAAATGGTCAATTAATCTGTATCCA 1680  
Db 1621 ATATATAAACCTCAACAATTTTCCAAACAATTTACCAAAATGGTCAATTAATCTGTATCCA 1680  
QY 1681 CAAAGGATTTCTGCATACATACACTTTAAACAATAATACCTAAATTTAGTGCATATTAA 1740  
Db 1681 CAAAGGATTTCTGCATACATACACTTTAAACAATAATACCTAAATTTAGTGCATATTAA 1740  
QY 1741 ACTTATTTGGTGGCATGACTATATCAACAGTTGCATGATATATGATACAAATTTATGTTA 1800  
Db 1741 ACTTATTTGGTGGCATGACTATATCAACAGTTGCATGATATATGATACAAATTTATGTTA 1800  
QY 1801 TTCTTTTCCATTTGCACCTGAAATACCATATATATAAAGAGAAATCCCATCATCAAAATTTGA 1860  
Db 1801 TTCTTTTCCATTTGCACCTGAAATACCATATATATAAAGAGAAATCCCATCATCAAAATTTGA 1860  
QY 1861 GCCTATATTGATTGATACCTCAGAAGAATCTGGCAGTGGAGGCTATATAAGGATTAAGCAA 1920  
Db 1861 GCCTATATTGATTGATACCTCAGAAGAACTGGCAGTGGAGGCTATATAAGGATTAAGCAA 1920  
QY 1921 TTGGGAAAGGATTTGGGAAGTTGGTAGTGAACATCTTCTCACCTGGAGCTCATGAGCAA 1980  
Db 1921 TTGGGAAAGGATTTGGGAAGTTGGTAGTGAACATCTTCTCACCTGGAGCTCATGAGCAA 1980  
QY 1981 CTTTGAATAGTTGAACCTGTGATGATATAGTATCTTAACACATTTTCCCTTGAATA 2040  
Db 1981 CTTTGAATAGTTGAACCTGTGATGATATAGTATCTTAACACATTTTCCCTTGAATA 2040  
QY 2041 GAAATTTGGCACAACAATTTTAAATTAATTTAGCAAAATTTTGGATATTTAAAGCTTCT 2100  
Db 2041 GAAATTTGGCACAACAATTTTAAATTAATTTAGCAAAATTTTGGATATTTAAAGCTTCT 2100  
QY 2101 TATAGAAAGAGATACCTGTATATTTAAGCCATGATGAGGTATATACAAATGTTAATATAT 2160  
Db 2101 TATAGAAAGAGATACCTGTATATTTAAGCCATGATGAGGTATATACAAATGTTAATATAT 2160  
QY 2161 TACTTGTACATGCAAAATTAATTTTATCATTTGGAGTCACTTTCTTTAAATTTAGT 2220  
Db 2161 TACTTGTACATGCAAAATTAATTTTATCATTTGGAGTCACTTTCTTTAAATTTAGT 2220  
QY 2221 AATGCCTTTTGGCTTAAATTTTCTCCTGATATTTAAATAGATACAGTAACCTTTCAATATG 2280  
Db 2221 AATGCCTTTTGGCTTAAATTTTCTCCTGATATTTAAATAGATACAGTAACCTTTCAATATG 2280  
QY 2281 TTAGTGTCTGTAATAATTTTTCATCTCTATTTTTCACCAATTTTATTCACATGTG 2340  
Db 2281 TTAGTGTCTGTAATAATTTTTCATCTCTATTTTTCACCAATTTTATTCACATGTG 2340  
QY 2341 CTCTTAATAAGTAGCATATAGTTAAATTTTAAAAATCCAAATATGCAATCACCTTTTAG 2400  
Db 2341 CTCTTAATAAGTAGCATATAGTTAAATTTTAAAAATCCAAATATGCAATCACCTTTTAG 2400  
QY 2401 GTTAAAAATTTAATCCATTTACATTTGTGCAAAATTCGACATATATATGTTCTTAAATCTA 2460  
Db 2401 GTTAAAAATTTAATCCATTTACATTTGTGCAAAATTCGACATATATATGTTCTTAAATCTA 2460  
QY 2461 TCATCTTACTAGTGGTTTCCATTTCTCTGCTCCAAATATTTTTCACAGCTTATAA 2520  
Db 2461 TCATCTTACTAGTGGTTTCCATTTCTCTGCTCCAAATATTTTTCACAGCTTATAA 2520  
QY 2521 CACAACCTTTTATAGAAAAGTTATACATAACACAGCATCAACTATTTTCAAGAACCCCAAT 2580  
Db 2521 CACAACCTTTTATAGAAAAGTTATACATAACACAGCATCAACTATTTTCAAGAACCCCAAT 2580  
QY 2581 AAGCAACAAAACCCAGACTAACAAAATGTGTAAACAGAAACTAATGACCTTTCTAAATC 2640  
Db 2581 AAGCAACAAAACCCAGACTAACAAAATGTGTAAACAGAAACTAATGACCTTTCTAAATC 2640  
QY 2641 AAACATTCATTTACTACATGCTATTTTACACACAGGGAAACCTCCATGGTTTACAGGC 2700  
Db 2641 AAACATTCATTTACTACATGCTATTTTACACACAGGGAAACCTCCATGGTTTACAGGC 2700

QY 2701 ATGTCATATTGAAATTAAGCTGCAATAGCTTTTATACAAATATATCCTCTCAAGAAAT 2760  
Db 2701 ATGTCATATTGAAATTAAGCTGCAATAGCTTTTATACAAATATATCCTCTCAAGAAAT 2760  
QY 2761 GAATCATTAAGACAGTAAATTAGGAGTTTCACAAATTTAAACATTTACGTAATTTTAAAT 2820  
Db 2761 GAATCATTAAGACAGTAAATTAGGAGTTTCACAAATTTAAACATTTACGTAATTTTAAAT 2820  
QY 2821 TATTGTCTTCAATTAATTTTAAATTAATTAAGCTGAGTTTCAAAAGTGATTTTTCCTCCAC 2880  
Db 2821 TATTGTCTTCAATTAATTTTAAATTAATTAAGCTGAGTTTCAAAAGTGATTTTTCCTCCAC 2880  
QY 2881 AAAGTGCCAACTTAAGCTTAGAGCTTTTCAAGTGTAACTTTTGGCCCTTAAAGTTAAGACA 2940  
Db 2881 AAAGTGCCAACTTAAGCTTAGAGCTTTTCAAGTGTAACTTTTGGCCCTTAAAGTTAAGACA 2940  
QY 2941 TATTCTGAGAAATATATAGTACATGATTTCTGATGCTATCTCTGTTAATAACAAA 3000  
Db 2941 TATTCTGAGAAATATATAGTACATGATTTCTGATGCTATCTCTGTTAATAACAAA 3000  
QY 3001 GATTTTCACATGATACCTATGTAACAAATCTCCATGTTCTACACATATATACCCAGAAC 3060  
Db 3001 GATTTTCACATGATACCTATGTAACAAATCTCCATGTTCTACACATATATACCCAGAAC 3060  
QY 3061 TTAAGTATATAATAATAAACAATAGCAAGCCCTTTAAAAAATAAAAAA 3114  
Db 3061 TTAAGTATATAATAATAAACAATAGCAAGCCCTTTAAAAAATAAAAAA 3114  
RESULT 2  
AAZ07531  
ID AAZ07531 standard; cdna: 850 BP.  
XX AAZ07531;  
AC AAZ07531;  
XX 26-NOV-1999 (first entry)  
DT Human SDCMP3 polypeptide encoding cdna.  
DE Binding compound; antibody binding site; primate; rodent; SDCMP3; SDCMP4;  
KW Schering dendritic cell membrane protein; dendritic cell physiology;  
KW genetic fingerprinting; cancer immunotherapy; abnormal proliferation;  
KW cancer; forensic; human; lectin 73; ss.  
XX Homo sapiens.  
OS  
XX Key Location/Qualifiers  
FH 108..596  
FT CDS /\*tag= a  
FT /product= "SDCMP3"  
XX WO9947673-A2.  
PN  
XX  
PD 23-SEP-1999.  
PF 16-MAR-1999; 99WO-US03740.  
XX  
PR 17-MAR-1998; 98US-0040111.  
XX (SCHE ) SCHERING CORP.  
XX Chalus L, Quan AB, Bates EEM, Gorman DM, Saeland S, Lebecque SJE;  
PI Philipps JH;  
XX  
DR WPI: 1999-562114/47.  
DR P-PSDB; AAY27446.  
XX  
PT Binding compound specific for primate or rodent Schering dendritic cell  
PT membrane proteins  
XX  
PS Claim 8; Page 80; 89pp; English.  
XX The invention relates to a binding compound comprising an antibody  
CC









PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;  
 XX WPI; 1999-591920/51.  
 XX New nucleic acid sequences expressed in ovarian, and some other, cancer  
 PT tissues, and derived polypeptides, for treatment of ovarian cancer and  
 PT identification of therapeutic agents.  
 XX  
 XX Claim 3; Page 206-207; 310pp; German.  
 XX This invention describes novel nucleic acid (cDNA) sequences (A) which  
 CC have anticancer activity and are highly expressed in ovarian tumor  
 CC tissue (and some also in testis and breast cancer tissue). The products  
 CC of the invention can be used for gene therapy. (A) are used (i) for  
 CC recombinant expression of polypeptides (B) and (ii) to isolate complete  
 CC genes. (B) are used (i) to identify agents suitable for treatment of  
 CC ovarian cancer; (ii) directly for treating this form of cancer  
 CC (including expression from gene therapy vectors) and (iii) for generation  
 CC of specific antibodies. (A) are identified by assembling ESTs (expressed  
 CC sequence tags) from a particular tissue type before comparison of  
 CC expression patterns. This allows a significantly longer fragment of the  
 CC gene to be revealed, so should reduce the number of failures associated  
 CC with the fact that ESTs from different libraries may represent different  
 CC parts of the same unknown gene, distorting the estimated frequency of  
 CC occurrence in a particular tissue. AA277450-277572 represent the human  
 CC ovarian tumor cDNA library derived EST fragments described in the method  
 CC of the invention and encode the protein fragments represented in  
 CC AA277505-Y76638.  
 XX  
 XX Sequence 3708 BP; 1036 A; 818 C; 671 G; 1183 T; 0 other;

Query Match 11.7%; Score 364.6; DB 20; Length 3708;  
 Best Local Similarity 88.3%; Pred. No. 1.3e-58;  
 Matches 462; Conservative 0; Mismatches 34; Indels 27; Gaps 5;  
 QY 2500 TATTTTACAGCTTATAACACAACTTTTATAGAAAAGTTATACATAACAGCATC 2559  
 DB 40 TTTTATTTTACAGCTTATAACACAACTTTTATAGAAAAGTTATACATAACAGCATC 99  
 QY 2560 AACATTTTTCAG- - - - - AACCCAAATAGCAACAAACAAACAGCACTAACAAATGT 2609  
 DB 100 AACATTTTTCAGAAACAATTAACCCGATAGCAACAAACAAACAGCACTAACAAATGT 159  
 QY 2610 GTAACAAGAACTAATGACCTTTCTAAATCAACAACTCAATATCTACAAATGCTATT 2669  
 DB 160 GTAACAAGAACTAATGACCTTTCTAAATCAACAACTCAATATCTACAAATGCTATT 219  
 QY 2670 ACAACAGGGAACCTCCATGTTTACAGGCATGTCATATTGAAATAAGCTCCATAG 2729  
 DB 220 ACAACAGGGAACCTCCATGTTTACAGGCATGTCATATTGAAATAAGCTCCATAG 279  
 QY 2730 C-TTTTATACAATTATCGCTCTCAAGAAAATGAATCAATTAAGCAGTAATTAGGAGTTC 2788  
 DB 280 CAATTTATACAATTATCGCTCTCAAGAAAATGAATCAATTAAGCAGTAATTAGGAGTTC 339  
 QY 2789 ACAATTTAAACATTTACGCTTAATTTTAAATTAATGCTTCAATATTTTAAATTTG 2848  
 DB 340 ACAATTTAAACATTTACGCTTAATTTTAAATTAATGCTTCAATATTTTAAATTTG 387  
 QY 2849 AAGTCTGAGTTTCAAAAGTGA- - - TTTTTCACCAAAAGGTGCCAACACTTTAAGCTAGAG 2905  
 DB 388 AAGTCTGAGTTTCAAAAGTGAITTTTTTTTCCCAACAAAGTTTCAACACTTTAAGCTAGAA 447  
 QY 2906 CTTTCAGTGTTAATCTTGGCCT- - - AAAAGTTAAGACATATTTCTGAGATCATTAATAGTCAC 2964  
 DB 448 CTTTCAGTGTTAATCTTGGCCT- - - AAAAGTTAAGACATATTTCTGAGATCATTAATAGTCAC 507  
 QY 2965 ATGATTTCTGATGCTATCTGCTCTGTTTAATAACAAAGATTTTCA 3007  
 DB 508 ATGATTTCTGATGCTATCTGCTCTGTTTAATAATAAAGCTTTTA 550

RESULT 7

AAC77686/C  
 ID AAC77686 standard; cDNA; 3748 BP.  
 XX  
 AC AAC77686;  
 XX  
 DT 08-FEB-2001 (first entry)  
 XX  
 DE Human cancer associated gene sequence SEQ ID NO:80.  
 XX  
 KW Human; cancer associated gene; cancer antigen; detection; cancer;  
 KW diagnosis; cytostatic; proliferative; vulnary; immunomodulator;  
 KW antidiabetic; antidiabetic; antirheumatic; antirheumatic; antirheumatic;  
 KW antidiabetic; antidiabetic; antirheumatic; antirheumatic; antirheumatic;  
 KW dermatological; neuroprotective; thrombolytic; coagulant; neoplastic;  
 KW vasotonic; antipsoriatic; antidiabetic; antidiabetic; antidiabetic;  
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;  
 KW allergic reaction; graft versus host disease; organ rejection;  
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;  
 KW neurological disease; drug screening; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200055350-A1.  
 XX  
 PD 21-SEP-2000.  
 XX  
 XX 08-MAR-2000; 2000WO-US05882.  
 PF  
 XX 12-MAR-1999; 99US-0124270.  
 PR  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 XX Rosen CA, Ruben SM;  
 PI  
 XX WPI; 2000-587533/55.  
 DR  
 DR P-PSDB; AAB43477.  
 XX  
 PT Novel isolated nucleic acids comprising sequences encoding peptides  
 PT useful for treating or diagnosing e.g. cancer -  
 XX  
 PS Claim 1; Page 671-673; 2352pp; English.  
 XX  
 CC AAC77607 to AAC78448 encode the human cancer associated proteins given  
 CC in AAB43398 to AAB44239. The proteins can have activities based on the  
 CC tissues and cells the genes are expressed in. Example of activities  
 CC include: cytostatic; proliferative; vulnary; immunomodulator;  
 CC antidiabetic; antidiabetic; antirheumatic; antirheumatic; antirheumatic;  
 CC antiinflammatory; antithyroid; antidiabetic; antidiabetic; antidiabetic;  
 CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;  
 CC neoplastic; vasotonic; antipsoriatic and antidiabetic. The  
 CC polynucleotides and polypeptides can be used for preventing, treating or  
 CC ameliorating medical conditions and diagnosing pathological conditions.  
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from  
 CC the present invention may be used to treat immune disorders by activating  
 CC or inhibiting the proliferation, differentiation or mobilisation of  
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune  
 CC disorders, allergic reactions, graft versus host disease and organ  
 CC rejection, modulate haemostatic or thrombolytic activity, modulate  
 CC inflammation, cancers, cardiovascular disorders, neurological disease and  
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,  
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to  
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of  
 CC the present invention.  
 XX  
 SQ Sequence 3748 BP; 1199 A; 685 C; 825 G; 1039 T; 0 other;  
 Query Match 11.7%; Score 364.6; DB 21; Length 3748;  
 Best Local Similarity 88.3%; Pred. No. 1.3e-58;  
 Matches 462; Conservative 0; Mismatches 34; Indels 27; Gaps 5;  
 QY 2500 TATTTTACAGCTTATAACACAACTTTTATAGAAAAGTTATACATAACAGCATC 2559  
 DB 3725 TTTTATTTTACAGCTTATAACACAACTTTTATAGAAAAGTTATACATAACAGCATC 3666





An alternative splice variant of this protein is also described (AAU000482). Other novel murine proteins include TANGO 210 (AAU000470) and INTERCEPT 400 (AAU00477). Six novel human proteins which include TANGO 210 (AAU00469), TANGO 364 (AAU00471), TANGO 366 (AAU00472), INTERCEPT 394 (AAU00473), INTERCEPT 400 (AAU00476) and TANGO 405 (AAU00479), and a rat INTERCEPT 400 (AAU00478) sequence are also described. The nucleic acids encoding these novel proteins are useful as modulating agents in regulating a variety of cellular processes and can be used to express the proteins in a host cell in gene therapy applications. Human and murine TANGO 405 proteins show sequence homology to murine Dectin-2. TANGO 405 modulates growth, proliferation, survival, differentiation, activity, morphology and movement/migration of human lymphocytes and bone marrow cells and tissues and can be used to prevent, diagnose or treat leukaemia, lymphomas and autoimmune disorders.



RESULT 13																									
ID AAV42554	standard; cDNA; 501 BP.																								
XX AC AAV42554;																									
XX DT 09-NOV-1998	(first entry)																								
XX DE Mouse	dectin-2 extracellular domain cDNA.																								
XX KW Dectin-2;	dendritic cell; c-type lectin; mouse; immunity; adjuvant; allergy; autoimmune disease; gene therapy; vaccine; diagnosis; drug screening; ss.																								
XX OS Mus sp.																									
XX FN W09828332-A2.																									
XX PD 02-JUL-1998.																									
XX PF 22-DEC-1997;	97WO-US23761.																								
XX PR 20-DEC-1996;	96US-0772440.																								
XX PA (TEXA ) UNIV TEXAS SYSTEM.																									
XX PI Ariizumi K,	Takashima A;																								
XX DR WPI; 1998-377594/32.																									
XX DR P-PSDB; AAW63022.																									
XX PT Nucleic acid encoding dendritic cell specific peptide(s) dectin-1 and -2 - useful, e.g. to regulate immune response, as vaccine adjuvants, for diagnosis and drug screening																									
XX PS Disclosure;	Page 153; 200pp; English.																								
XX CC This CDNA sequence codes for the extracellular domain (see AAW63022) of a novel murine dendritic cell (DC) specific member of the C-type lectin family, termed DC-associated C-type lectin-2, or dectin-2 (see also AAW63010), that is essential for DC-mediated T cell activation. Full-length cDNA (see AAV42549) was isolated by subtractive hybridisation from BALB/c mouse epidermis DC line XS52. The invention provides: dectin-1 and -2 polypeptides (see also AAW63009-22 and AAW69236-37), useful for purifying T cells, for detecting autoantibodies and for up-regulating immunity e.g. as vaccine adjuvants; dectin DNA (see AAV42548-54, AAV42558-60 and AAV44850-51); expression vectors; recombinant host cells; probes and primers; antibodies; compounds that modulate dectin-mediated activation of T cells; transgenic animals; and dectin ligands. Dectin expression can also be down-regulated to treat allergy and autoimmune disease. Recombinant extracellular domain has been expressed as a His-tagged protein (see AAW69237) useful for antibody production.																									
XX SQ Sequence	501 BP; 142 A; 108 C; 115 G; 136 T; 0 other:																								
<table border="0"> <tr><td>Query Match</td><td>9.7%;</td><td>Score</td><td>303.2;</td><td>DB 19;</td><td>Length</td><td>501;</td></tr> <tr><td>Best Local Similarity</td><td>78.4%;</td><td>Pred. No.</td><td>2.4e-47;</td><td></td><td></td><td></td></tr> <tr><td>Matches</td><td>378;</td><td>Conservative</td><td>0;</td><td>Mismatches</td><td>98;</td><td>Indels</td><td>6;</td><td>Gaps</td><td></td></tr> </table>		Query Match	9.7%;	Score	303.2;	DB 19;	Length	501;	Best Local Similarity	78.4%;	Pred. No.	2.4e-47;				Matches	378;	Conservative	0;	Mismatches	98;	Indels	6;	Gaps	
Query Match	9.7%;	Score	303.2;	DB 19;	Length	501;																			
Best Local Similarity	78.4%;	Pred. No.	2.4e-47;																						
Matches	378;	Conservative	0;	Mismatches	98;	Indels	6;	Gaps																	
QY	305	GC AAAAGGCTGTCTGAACACTCAGTATCATTCAGTGCTCACCTGGTTCAAGTGAAGGGA	364																						
Dd	20	GTAGAAGACTATATGAACCTCCACATACCATTCCAGTCTCACCTGGCTTCAGTGAAGGGA	79																						
QY	365	CAAAGTGCCAG-----CCTGGGGATGTTGCCAGCTTCTTTGAAGTCATTTGGTTCCA	418																						
Dd	80	CTATGGTCTCAGAAAAAATGTGGGGATGCTGCCAAATCACTGGNAGTCATTTGGCTTCCA	139																						
QY	419	GTTCGTACTTCATTTCCAGTGAAGAAGGTTTGGTCTTAAGAGTCAGCACAACCTGTGTTG	478																						





QY	223	GTGGGATTTCCATTGGCACTCCCTACGTGCTTGCTTCATTGTGAGCTGTGTAGTAACATTAC	282
Db	88	GCAGTCGATCCATCTTGCTCCCTCAGTGTCTGTTTCACGTGTGAGTCTGTGGTGGCCTCAC	147
QY	283	CATTTTACATATGGTGAACCTGCCAAAGCGTCTGAACATA-----CAGTCATAT	333
Db	148	AAATTTATGTATAGCAAACTGTCAAGAGCGTGTCCAAGTTACGAGAGTATCAACAGTAT	207
QY	334	CATTCAAGTCTCAGCTGTTTCAGTGAAGGACAAAGTGCACGCGCTGGGATGTTCGCCCA	393
Db	208	CATCCAGCCTGACCTGGTCTATGGGAAGAAAGGACATAGAGATTGGAGCTGCTGCCCA	267
QY	394	GCTTCTTGGAACTCATTTGGTTCAGTGTGCTACTTTCATTTCCAGTGAAGAGAAGTTTGG	453
Db	268	ACCCCTTGGACTTCATTTGACTAGTGTGCTACTTTTACTCTGGGATGCAATCTTGG	327
QY	454	TCTAAGAGTGAAGCAACTGTGTGAGATGGGAGGACACATTTTGGTGTCTTCAACACAGAA	513
Db	328	ACTAAGAGTCAAAAGAACTGTCTGTGATGGGGGCTGATCTGGTGGTATCAACACACAGG	387
QY	514	GCAGAGCAGAAATTCATTTGTCAGCAGCTGAATGAGTCATTTCTCTATTTTCTGGGGCTT	573
Db	388	GAAGACAGGATTCATCATTCAGATCTGAAAAGAAATTCCTCTTATTTCTGGGGCTG	447
QY	574	TCAGACCCACAAGTGAATAAATTTGGCAATGGATGATGAAGACACACCTTATGAGAAAAT	633
Db	448	TCAGATCCAGGGGGTCGGCGACATTTGGCAATGGTGTGACCAACACCATACAATGAAAT	507
QY	634	GTCAAGATTTGGCACCCTAGGTGAGCCCAATCATTTCTGACAGAGCAATGTGCTTCAATAGTC	693
Db	508	GTCAATTTTGGCACTCAGGTGAACCCAAATACCTTTGATGAGCGTGTGGCGAATAAAT	567
QY	694	TTCTGAAACCTACAGGATGGGCTGGAATGATGTTATCTGTGAAACTAGGAAGGAATCA	753
Db	568	TTCCGTCTCTCAGAGAAATGGGCTGGAATGACATTCAGCTGTCTATGATACCTCAGAACTCA	627
QY	754	ATATGTGAGATGAATAAGATTTTACCTATGAGTGAAGCTT	793
Db	628	ATTTCGAAGATGAAGAAGATCTCATATAAATGAATATT	667

Search completed: February 17, 2003, 09:02:36  
Job time : 696.327 secs



QY 100 AGGAGAGAGGTACAAAAGGTTCTGGACCTTCTCAACACAGGGAGCCTGCATAATGATG 159  
Db 98 --GGACAGATGCAAGAAACCCCT--GACCTTCTGAACATA---CACCTCAACAATGGTG 151  
QY 160 CAAGAGCAGCAACCTCAAAAGTACAGAGAAAGAGGCTGTTGTCCTCAGACTCTGGTCT 219  
Db 152 CAGGAAGACAATCCA-----AGGAGGAGGTCTCTGGACCTCAGACTCTGGTCA 205  
QY 220 GTGGCTGGGATTTCATTTGCATCTCCTCAGTGTCTTTCATTTGAGCTGTGTAGTAAT 279  
Db 206 GCTGCTGTGATTTCATTTGCTTGTAGTACCTGCTTTCATTTGCGAGCTGTGTGTGACT 265  
QY 280 TACCATTTCATATATGGTGAACACTGGCAAAAGGCTGTCTGAACACACACTCATATCAATCA 339  
Db 266 TACCAATTTATTATGACACCGCCAGTAGAGACTATATGAACATTCACACATACCATCC 325  
QY 340 AGTCTCACCTGCTTCAGTGAAGGACAAAGGTGCCAG-----CCTGGGATGTTGCCCA 393  
Db 326 AGTCTCACCTGCTTCAGTGAAGGACTATGTGTCAGAAAAAATGCGGGATGCTGCCCA 385  
QY 394 GCTTCTTGAAGTCAATTTGGTTCAGTGTCTTCTACTTTCATTTCCAGTGAAGAAAGTTTG 453  
Db 386 AATCACTGGAAGTCAATTTGGCTCCAGCTGCTACCTCAATTTCTACCAAGGAGAACTTCTGG 445  
QY 454 TCTAAGAGTGACAGAACTGTCTGAGATGGGAGCACATTTGGTGTCTTCAACACAGAA 513  
Db 446 AGCAGAGTGACAGAACTGTCTGAGATGGGAGTGGGCTCACTGGTGTGATCAATCACTGAA 505  
QY 514 GCAGAGCAGAAATTTCAATTTGTCAGCAGCTGAATGAGTCAATTTCTTATTTCTGGGGCTT 573  
Db 506 GCGGAGCAGAAATTTCAATCCAGCAGCTGAATGAGTCAATTTCTTACTTCTGGGCTT 565  
QY 574 TCAGACCCACAAGGTAATAATTTGCAATTTGCAATTTGATGATGAACACACCTTATGAGAAAAAT 633  
Db 566 TCGGATCCACAAGGTAATAATTTGCAATTTGCAATTTGATGATGAACACACCTTATGAGAAAAAT 625  
QY 634 GTCAGATTTTGCACCTAGGTGAGCCCAATCATTTCTGCAGAGCAATGCTTCAATAGTC 693  
Db 626 GTCAGTTCTGCGACCCCAATGAGCCCAATCTTCCAGAGAGCGGTGTTTCAATAGTT 685  
QY 694 TTCTGGAACCTACAGGATGGGCTGGAATGATGTTATCTGTAACCTAGAGAAATCA 753  
Db 686 TACTGGAATCCCTCGAAATGGGCTGGAATGATGTTTCTGTGATAGTAACACAATCA 745  
QY 754 ATATGTGAGTCAATAAGATTTACCTATGAT 785  
Db 746 ATATGTGAAATGAAGAATTTACCTATGAT 777

RESULT 2

US-08-772-440-20  
; Sequence 20, Application US/08772440  
; Patent No. 6046158  
; GENERAL INFORMATION:  
; APPLICANT: Ariizumi, Kiyoshi  
; APPLICANT: Takashima, Akira  
; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE  
; TITLE OF INVENTION: LECTINS, DECTIN-1 AND DECTIN-2; COMPOSITIONS AND USES  
; NUMBER OF INVENTION: 42  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/772,440  
; FILING DATE: CONCURRENTLY HERewith  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Parker, David L.  
; REGISTRATION NUMBER: 32,165  
; REFERENCE/DOCKET NUMBER: UTDX:493  
; TELEPHONE: 512/418-3000  
; TELEFAX: 512/474-7577  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 501 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-772-440-20  
  
Query Match 9.7%; Score 303.2; DB 3; Length 501;  
Best Local Similarity 78.4%; Pred. No. 7.8e-58;  
Matches 378; Conservative 0; Mismatches 98; Indels 6; Gaps 1;  
  
QY 305 GCAAAAGGCTGCTGAACCTACACTCATATCACTCAAGTCTCACTGCTTCAGTGAAGGGA 364  
Db 20 GTGAAGACTATATGAACCTTCACACATACCATTCACCTCTCACTGCTTCAGTGAAGGGA 79  
QY 365 CAAAGGTGCCAG-----CCTGGGATGTTGCCAGCTTCTTGGAAAGTCAATTTGTTTCCA 418  
Db 80 CTATGGTGTGCAAAAAAATGCGGGATGCTGCCCAATCACTGGAAGTCAATTTGGTCCA 139  
QY 419 GTTGTCTACTTCAATTCAGTGAAGAGAGGTTTGGTCTTAAGAGTGAGCAGAACTGTGTTG 478  
Db 140 GCTGCTACCTCAATTTCTACCAAGGAGAACTTCTGGAGCACCACTGAGCAGAACTGTGTT 199  
QY 479 AGATGGGAGCACATTTGGTGTCTCAACACAGACAGCAGAGAAATTTCAATTTGTTCCAG 538  
Db 200 AGATGGGAGTCACTGTTGGTGTATCAATCACTGAACGGAGCAGAAATTTCAATCACCAGC 259  
QY 539 AGCTGAATGAGTCAATTTCTTATTTCTGGGCTTTTCAGACCCACAGTAATAATAAT 598  
Db 260 AGCTGAATGAGTCACTTCTTACTTCTGGTCTTTCAGTCAAAATGTCAGGTTCTGGCACCCCATGAAC 319  
QY 599 GCAATGGATTGTAAGACACCTTATGAGAAAAATGTCAGAAATTTGGCACCTAGGTGAGC 658  
Db 320 GCAATGGATTGTAAGACACCTTATGAGAAAAATGTCAGTCAAAATGTCAGGTTCTGGCACCCCATGAAC 379  
QY 659 CCAATCACTTCGACAGCAATGCTTCAATAGTCTTCTGGAACCTCAGAGTGGGCT 718  
Db 380 CCAATCTTCAGAGAGCGGTGTTTCAATAGTCTTCTGGAATCCTTCGAATGGGCT 439  
QY 719 GGAATGATGTTATCTGTGAACCTAGAGAAATTCATATGTGAGATGAATAAGATTTACC 778  
Db 440 GGAATGATGTTTCTGTGATAGTAAACACAATTCATATGTGAAATGAAGAAGATTTACC 499  
  
QY 779 TA 780  
Db 500 TA 501

RESULT 3

US-08-772-440-22  
; Sequence 22, Application US/08772440  
; Patent No. 6046158  
; GENERAL INFORMATION:  
; APPLICANT: Ariizumi, Kiyoshi  
; APPLICANT: Takashima, Akira  
; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE  
; TITLE OF INVENTION: LECTINS, DECTIN-1 AND DECTIN-2; COMPOSITIONS AND USES  
; NUMBER OF INVENTION: 42  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P.O. Box 4433

CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/772,440  
FILING DATE: CONCURRENTLY HERewith  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Parker, David L.  
REGISTRATION NUMBER: 32,165  
REFERENCE/DOCKET NUMBER: UXD:493  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 393 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-772-440-22

Query Match 8.4%; Score 260.2; DB 3; Length 393;  
Best Local Similarity 78.9%; Pred. No. 2e-48;  
Matches 310; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 388 TGCCACGCTCTTGGAGTCATTTGGTCCAGTGTCTACTTTCATTTCCAGTGAAGAG 447  
DB 1 TGCCCAAACTCACTGGAAGTCATTTGGCTCCAGCTGTACCTCATTTCTACCAAGGAGAC 60  
QY 448 GTTTGGTCTAAGAGTGAGCAGAACTGTGTTGAGATGGGAGCACATTTGGTTGTTCAC 507  
DB 61 TTCTGGAGCACCAGTGAGCAGAACTGTGTTGAGATGGGAGTCATCTGCGGTGATCAAT 120  
QY 508 ACAGAGCAGCAGAGAAATTTTCATTTGCCAGCAGTGAATGAGTCATTTCTTTATTTCTG 567  
DB 121 ACTGAAGCGGAGCAGAAATTTTCATCACCCAGCAGCTGAATGAGTCACCTTTCTTACTTCTG 180  
QY 568 GGGCTTTTCAGACCCACAGAGTAATTAATGCGCAATGGAATGATTAAGACACCTTATGAG 627  
DB 181 GGTCTTTTCGAGTCCCAAGGTAATGCGCAATGGAATGGAATGGAATGATATCTCTTTCAGT 240  
QY 628 AAAAATGTCAGATTTTGGCACCTAGGTGAGGCCCAATCATTTCTGCAGACGAATGTGCTTCA 687  
DB 241 CAAAATGTCAGTTCTGCGACCCCATGAAACCCCAATCTTCCAGAGAGCGGTGTGTTCA 300  
QY 688 ATAGTCTTCTGGAACCTACAGGATGGGCTGGAATGATGTTATCTGTGAAACTAGAAGG 747  
DB 301 ATAGTTTACTGGAATCTCTGGAATGGGCTGGAATGATGTTTCTGTGTAGTAAACAC 360  
QY 748 AATTCAATATGTGAGATGAATAAGATTACCTA 780  
DB 361 AATTCAATATGTGAATGAAGAAGATTACCTA 393

RESULT 4  
US-09-111-470-1  
Sequence 1, Application US/09111470  
Patent No. 6277959  
GENERAL INFORMATION:  
APPLICANT: Valladeau, Jenny  
APPLICANT: Ravel, Odile  
APPLICANT: Bates, Elizabeth E.M.  
APPLICANT: Ford, John  
APPLICANT: Saeland, Sem  
APPLICANT: Lebecque, Serge J.E.  
TITLE OF INVENTION: Mammalian Membrane Protein Genes;

TITLE OF INVENTION: Related Reagents  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DNAX Research Institute  
STREET: 901 California Avenue  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304-1104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/111,470  
FILING DATE: 08-JUL-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/053,080  
FILING DATE: 09-JUL-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ching, Edwin P.  
REGISTRATION NUMBER: 34,090  
REFERENCE/DOCKET NUMBER: SF0695  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650)852-9196  
TELEFAX: (650)496-1200  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1104 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 242..952  
US-09-111-470-1

Query Match 6.3%; Score 197; DB 4; Length 1104;  
Best Local Similarity 67.8%; Pred. No. 2e-34;  
Matches 291; Conservative 0; Mismatches 135; Indels 3; Gaps 1;

QY 368 AGTGCCACCCCTGGGGATGTTGCCAGCTTCTTGAAGTCATTTGGTTCACAGTGTCTACT 427  
DB 537 AAGAGACAGCCCTGGAGCTGTGCCCAAGAAATTTGAAGTCATTTAGTTCACACTGCTACT 596  
QY 428 TCATTTCACAGTGAAGAGAGGTTTGGTCTTAAGAGTGAGCAGAACTGTGTTGAGATGGGAG 487  
DB 597 TTATTCTACTCAATCAGCATCTTGCAAGACAGTGAAGGACTGTGCTAGAATGGAGG 656  
QY 488 CACATTGTTGTTGTTCAACAGACAGAGCAGAAATTTTCATTTGCCAGCAGCTGAATG 547  
DB 657 CTCACCTGCTGTGATAAAACACTCAAGAAAGACAGGATTTTCATCTTCCAGAATCTGCAAG 716  
QY 548 AGTCATTTCTTATTTTCTGCGGCTTTCAGACCCCAAGGTAATTAATTTGCAATGGA 607  
DB 717 AAGATCTGCTTATTTTGTGGGCTCTCAGATCCAGAGGTCAGGCACATTTGGCAATGGG 776  
QY 608 TTGATAAGACACCTTATGAGAAAAATGTCAGATTTTGGCACCTAGGTGAGGCCCAATCAT 667  
DB 777 TTGTCAGACACCATACATAATGAAGTTCCACATTTCTGGCATCCAGTGAAGCCAGTGATC 836  
QY 668 CTGCAGACGAATGTGCTTCAATAGTCTTCTCTGGAA --- ACCTACAGATGGGCTGGAATG 724  
DB 837 CCAATGAGCGCTCGCTTGTGCTAAATTTTCGTAATAACCCCAAGATGGGCTGGAATG 896  
QY 725 ATGTTATCTGTAAACTAGAAGGAATTCATATGTGAGATGAATTAAGATTTACCTATGAG 784  
DB 897 ATGTTAAATGTTCTGTGCTCCTCAAGGTCAGTTTGTGAGATGATGAAGATCCACTTATGAA 956  
QY 785 TAGAAGCTT 793





db 61 CAAGGTAATGGCAAAATGGCAATGGATCGATGATCTCCTTTCAGTCAAAATGTCAG 116

Qy	2132	TGATGAGGTATACAAATTATAATTTACTTGTACATGCCAATAATTTTTTATC	2191
Db	28692	TTATTGGGTACATAACGGTTTAGAATGTTATGTCCTTGATGAATGCACCTTTTCATC	28751
Qy	2192	ATTGTGGAGTCACTTCTCTTAATAATTAGTAATGCCCTTGGCTTTAATTT-----TTCTCCT	2247











GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 17, 2003, 08:49:16 ; Search time 114.038 Seconds  
(without alignments)  
13907.882 Million cell updates/sec

Title: US-09-766-511b-51  
Perfect score: 3114  
Sequence: 1 cttaatgttggaagtctctt.....tttaaaaaaaaaaaaaaaaaa 3114

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 424239 seqs, 254661826 residues

Total number of hits satisfying chosen parameters: 848478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published-Applications\_NA.\*

```
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
C 1	364.6	11.7	3748	10	US-09-925-301-80
C 2	360.6	11.6	693	10	US-09-833-381-994
C 3	311.6	10.0	528	10	US-09-833-381-995
4	274.8	8.8	642	12	US-10-090-466-1
5	232.4	7.5	549	12	US-10-090-466-3
6	197	6.3	1091	10	US-09-764-870-199
7	197	6.3	1096	10	US-09-764-870-38
8	197	6.3	1104	9	US-09-862-802-1
9	185.2	5.9	1036	10	US-09-833-381-119
C 10	151.2	4.9	291	10	US-09-833-381-996
C 11	137.4	4.4	1418	9	US-09-862-802-7
12	120.6	3.9	758	10	US-09-833-381-1339
13	117.4	3.8	968	9	US-09-965-529-40
14	117.4	3.8	997	9	US-09-905-291A-376
15	117.4	3.8	997	9	US-09-902-853-376
16	117.4	3.8	997	9	US-09-907-824-376
17	117.4	3.8	997	9	US-09-907-841-376
18	117.4	3.8	997	9	US-09-904-011-376
19	117.4	3.8	997	9	US-10-174-590-23

20	117.4	3.8	997	9	US-10-176-758-23	Sequence 23, Appl
21	117.4	3.8	997	9	US-10-175-737-23	Sequence 23, Appl
22	117.4	3.8	997	9	US-09-906-742-376	Sequence 376, App
23	117.4	3.8	997	9	US-10-173-706-23	Sequence 23, Appl
24	117.4	3.8	997	9	US-10-175-738-23	Sequence 23, Appl
25	117.4	3.8	997	9	US-10-175-752-23	Sequence 23, Appl
26	117.4	3.8	997	9	US-10-176-482-23	Sequence 23, Appl
27	117.4	3.8	997	9	US-10-176-757-23	Sequence 23, Appl
28	117.4	3.8	997	9	US-10-176-913-23	Sequence 23, Appl
29	117.4	3.8	997	9	US-10-180-552-23	Sequence 23, Appl
30	117.4	3.8	997	9	US-10-180-557-23	Sequence 23, Appl
31	117.4	3.8	997	9	US-09-906-838-376	Sequence 376, App
32	117.4	3.8	997	9	US-09-907-613-376	Sequence 376, App
33	117.4	3.8	997	9	US-09-907-942-376	Sequence 376, App
34	117.4	3.8	997	9	US-10-173-700-23	Sequence 23, Appl
35	117.4	3.8	997	9	US-10-174-572-23	Sequence 23, Appl
36	117.4	3.8	997	9	US-10-174-579-23	Sequence 23, Appl
37	117.4	3.8	997	9	US-10-174-582-23	Sequence 23, Appl
38	117.4	3.8	997	9	US-10-174-588-23	Sequence 23, Appl
39	117.4	3.8	997	9	US-10-175-739-23	Sequence 23, Appl
40	117.4	3.8	997	9	US-10-175-740-23	Sequence 23, Appl
41	117.4	3.8	997	9	US-10-175-743-23	Sequence 23, Appl
42	117.4	3.8	997	9	US-10-176-488-23	Sequence 23, Appl
43	117.4	3.8	997	9	US-10-176-492-23	Sequence 23, Appl
44	117.4	3.8	997	9	US-10-176-747-23	Sequence 23, Appl
45	117.4	3.8	997	9	US-10-176-750-23	Sequence 23, Appl

## ALIGNMENTS

```
RESULT 1
US-09-925-301-80/c
; Sequence 80, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 80
; LENGTH: 3748
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-301-80
```

Query Match	11.7%	Score 364.6;	DB 10;	Length 3748;
Best Local Similarity	88.3%;	Pred. No. 7.8e-59;		
Matches 462;	Conservative 0;	Mismatches 34;	Indels 27;	Gaps 5;
QY 2500	TATTTTTTTTACAGCTTATAACACAACTTTTATTAGAAAAGTTATACATACACAGCATC	2559		
Db 3725	TTTTTTTTTTTACAGCTTATAACAACTTTTATTAGAAAAGTTATACATACAGCATC	3666		
QY 2560	AACTATTTTCAAG-----AACCCAATAAGCAACAAACCAAGCACTAACAAATGT	2609		
Db 3665	AACTATTTTCAAGACAATATTAAACCCGATAGCAACAAACCAAGCACTAACAAATGT	3606		
QY 2610	GTAAACAAGAACTAATACCTTTCTTAAATCAAACTTCAATTTATCAATGTCTATTT	2669		
Db 3605	GTAAACAAGAACTAATACCTTTCTTAAATCAAACTTCAATTTATCAATGTCTATTT	3546		
QY 2670	ACAAACAGGGAACCTCCATGGTTTACAGGATGTCATATTGAAATAAAGCTGCAATAG	2729		
Db 3545	ACAAACGGGGAAACCTCCTTGGTTTACAGGCACATCATATTGAATATAAGCTGCAATAG	3486		

```
QY 2730 C-TTTTATACAAATTATCGCTCTCAAGAAAATGAATCATTTAAGACAGTAATTAGAGGTC 2788
Db 3485 CAATTTTATACAAATTACCACCTCTGAAGAACTGAATCATTTAAACAGCTAATACGAGTC 3426
QY 2789 ACAAAATTTAAACATTTTCAGCTAATTTTAAATATTATGCTTCATATATTTTAAATTTATG 2848
Db 3425 ACAAAATTTAAACATTTTCACATAATTTTAAATATTATGSGT-----ATACACTG 3378
QY 2849 AAGTCGTAGTTTCAAAAGTGA---TTTTTCCCAAAAGTGCCAACTTTAAAGCTAGAG 2905
Db 3377 AAGTCGTAGTTTCAAAAGTGATTTTTCCTCCCAAAAGTTTCAACACTTAAAGCTAGAA 3318
QY 2906 CTTTCAGTGTAACTTTGCCCT-AAAAGTTAAGACATATTCGAGAAATCAATAATAGTCA 2964
Db 3317 CTTTCAGTGTAACTTTGCCCTTAAAGCTTAAAGACATATTCGATATATCAATAACAGTCA 3258
QY 2965 ATGATTTCTGATGCTATCTGCTCTGTTTAAACAAAGATTTCA 3007
Db 3257 ATGATTTCTGATGCTATCTGCTCTGTTTAAATAAAGTCTTTA 3215

RESULT 2
US-09-833-381-994/c
; Sequence 994, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 994
; LENGTH: 693
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(693)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-994

Query Match 11.6%; Score 360.6; DB 10; Length 693;
Best Local Similarity 76.7%; Pred. No. 2.7e-58;
Matches 468; Conservative 0; Mismatches 135; Indels 7; Gaps 2;

QY 183 AGAGAAAAGAGGCTGGTGTCTCCCTGAGACTCTGGTCTGTGGCTGGGATTTCCATTTGCACT 242
Db 690 AGGGAAGGAGTCTGCTGGACCTGAGACTCTGGTCTGGTGTGATTTCCATGTTACT 631
QY 243 CCTCAGTGTCTTCATTTGAGCTGTGTAGTACTTACCAATTTACATATGTTGAAC 302
Db 630 CTTGAGTACTGCTTTTCATTTGCGAGCTGTGTGTGACTTACCAATTTATATGACACGCC 571
QY 303 TGSCAAAAGGCTGTCTGAACTACACTCATATCATTTCAAGTCTCAGCTTCCAGTGAAG 362
Db 570 CAGTAGAAGACTATATGAATCTTACACATACCACTTCCAGTCTCAGCTGCTTCCAGTGAAG 511
QY 363 GACAAAAGTGCCAG-----CCTGGGGATGTTGCCAGCTTCTTGGAAAGTCAATTTGGTTC 416
Db 510 GACTATGGTGTACAGAAAAATGTGGGATGCTGCCCAAACTACTGGAAGTCAATTTGGCTC 451
QY 417 CAGTTGCTACTTCAATTTCCAGTGAAG-AGAAGGTTTGGTCTAAGAGTGAAGCAACTGTG 475
Db 450 CAGCTGCTACTTCAATTTCTACCAAGGAAGAACTTCTGGAGCACCAGTGAAGCAACTGTG 391
QY 476 TTCAGATGGGACACATTTGGTGTCTTCAACACAGACGAGCAGAAATTTTCATTGTCC 535
Db 390 TTCATATGGGGGCTCATCTGGTGGTATCATACTAGAACCGGAGCAAGATTTTCATCACCC 331
```

```
QY 536 AGCAGCTGAATGAGTGCATTTTCTTATTTTCTGGGCTTTTCAGACCCACACAGCTAATAATA 595
Db 330 AGCAGCTGAATGAGTGCATTTTCTTACTTCTCGGTCTTTTCGGATCCACAAGGTAATGGCA 271
QY 596 ATTGGCAATGGATTGTAAGACACACCTTTATGAGAAAAATGTCAGATTTTGGCACCTAGGTG 655
Db 270 AATGGCAATGGATTGTAAGTATGATCTCTTTCAGTCAAAATGTCAGGTCTTGGCACCCCATG 211
QY 656 AGCCCAATCATTTCTGCAGAGCAATGTGCTTCAATAGTCTTCTTGGAAACCTACAGGATGG 715
Db 210 AACCAATCTTCCAGAAGAGCGGTGTGTTTCAATAGTCTTCTTGGAAATCCTTCCAAATGGG 151
QY 716 GCTGGAATGATGTTATCTGTGAAACTAGAGGAATTTCAATATGTCAGATCAATAGATTT 775
Db 150 GCTGGAATGATGTTTCTGTGTAGTAGTAAACACAATTTCAATATGTTGAAATGAAGAAGATTT 91
QY 776 ACCTATGAGT 785
Db 90 ACCTATGAGT 81

RESULT 3
US-09-833-381-995/c
; Sequence 995, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 995
; LENGTH: 528
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(528)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-995

Query Match 10.0%; Score 311.6; DB 10; Length 528;
Best Local Similarity 76.0%; Pred. No. 3.1e-49;
Matches 399; Conservative 0; Mismatches 120; Indels 6; Gaps 1;

QY 183 AGAGAAAAGAGGCTGGTGTCTCCCTGAGACTCTGGTCTGTGGCTGGGATTTCCATTTGCACT 242
Db 525 AGGGAAGGAGTCTGCTGGACCTGAGACTCTGGTCTGAGCTGTGTGATTTCCATGTTACT 466
QY 243 CCTCAGTGTCTTCTTCATTTGAGCTGTGTAGTAACTTACCAATTTACATATGTTGTAAC 302
Db 465 CTTGAGTACTGTTTTCATTTCCGAGCTGTGTGTGAGCTTACCAATTTATATGACACGCC 406
QY 303 TGGCAAAAAGGCTGTCTGAACTACACTCATATCATTTCAAGTCTCACCTGCTTCAGTGAAG 362
Db 405 CAGTAGAAGACTATATGAATCTTACACATACCACTTCCAGTCTCACCTGCTTTCAGTGAAG 346
QY 363 GACAAAAGTGCCAG-----CCTGGGGATGTTGCCAGCTTCTTGGAAAGTCAATTTGGTTC 416
Db 345 GACTATGGTGTACAGAAAAATGTGGGATGCTGCCAAATCACTTGGAAAGTCAATTTGGCTC 286
QY 417 CAGTTGCTACTTCAATTTCCAGTGAAGAGAGGTTTGGTCTTAAGAGTGAAGCAACTGTGT 476
Db 285 CAGCTGCTACTTCAATTTCTACCAAGGAGAACTTCTGGAGCACCAGTGAAGCAACTGTGT 226
QY 477 TGAGATGGGACACATTTGGTGTCTTCAACACAGAGCAGACAGAAATTTTCATTTGTC 536
Db 225 TCAGATGGGGCTCATCTGGTGGTGTATCAATACTAGAACGGAGCAGAAATTTTCATCACCCA 166
```



```
RESULT 6
US-09-764-870-199
; Sequence 199, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT214
; CURRENT APPLICATION NUMBER: US/09/764,870
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 199
; LENGTH: 1091
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-870-199

Query Match      6.3%; Score 197; DB 10; Length 1091;
Best Local Similarity 67.8%; Pred. No. 7.1e-28;
Matches 291; Conservative 0; Mismatches 135; Indels 3; Gaps 1;

QY 368 AGGTGCCAGCCTGGGGATGTTGCCAGCTTCTTGGAAAGTCATTTGGTTCCAGTTGCTACT 427
Db 314 AAGAGACAGCCTGGAGCTGTTGCCAAAGAAATTGGAAGTCATTTAGTTCCAACTGCTACT 373

QY 428 TCATTTCCAGTGAAGAGAGGTTTGGTCTAAGAGTCAGCAGAACTCTGTGTGAGATGGAG 487
Db 374 TTATTTCTACTGAATCAGCATCTTGGCAAGACAGTCAGAAAGACTGTGCTAGAAATGGAG 433

QY 488 CACATTTGGTTGTTCAACAGACAGCAGAGAAATTTCAATTCGCCAGCAGTGAATG 547
Db 434 CTCACCTGCTGGTGATAACACTCAAGAGACAGAGGATTTCACTTCCAGAACTGCGAAG 493

QY 548 AGTCATTTTCTTATTTCTGGGGCTTTCAGACCCACAGGTAATAATAATTTGGCAATGGA 607
Db 494 AAGAATCTGCTTATTTGTGGGGCTCTCAGATCCAGAAAGGTCAGCGACATTTGGCAATGG 553

QY 608 TTGATAGACACCTTATGAGAAATGTCAGATTTTGGCACTTCCAGCAGCTGTAATG 567
Db 554 TTGATCAGACACATACATAAGTAAAGTTCCACATTTGGCATCCAGTCAGCCAGCTGATC 613

QY 668 CTCAGACGAATGTCTTCAATAGTCTTCTGGAA---ACCTACAGGATGGGCTGGAATG 724
Db 614 CCAATGAGCGCTGCTGTGCTAAATTTTCGTAATCACCACAAAGATGGGCTGGAATG 673

QY 725 ATGTTATCTGTGAACACTAGAAGAAATTCATATGTGAGATGAATAAGATTTACCTATGAG 784
Db 674 ATGTTAATTGCTTGTGCTCCTCAAAGGTCAGTTTGTGAGATGATGAATCCACTTATGAA 733

QY 785 TAGAAGCTT 793
Db 734 CTGAACAT 742

RESULT 7
US-09-764-870-38
; Sequence 38, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT214
; CURRENT APPLICATION NUMBER: US/09/764,870
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
; LENGTH: 1096
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-870-38

Query Match      6.3%; Score 197; DB 9; Length 1104;
Best Local Similarity 67.8%; Pred. No. 7.2e-28;
Matches 291; Conservative 0; Mismatches 135; Indels 3; Gaps 1;

QY 368 AGGTGCCAGCCTGGGGATGTTGCCAGCTTCTTGGAAAGTCATTTGGTTCCAGTTGCTACT 427
Db 537 AAGAGACAGCCTGGAGCTGTTGCCAAAGAAATTGGAAGTCATTTAGTTCCAACTGCTACT 596
```

```
; ORGANISM: Homo sapiens
US-09-764-870-38

Query Match      6.3%; Score 197; DB 10; Length 1096;
Best Local Similarity 67.8%; Pred. No. 7.2e-28;
Matches 291; Conservative 0; Mismatches 135; Indels 3; Gaps 1;

QY 368 AGGTGCCAGCCTGGGGATGTTGCCAGCTTCTTGGAAAGTCATTTGGTTCCAGTTGCTACT 427
Db 324 AAGAGACAGCCTGGAGCTGTTGCCAAAGAAATTGGAAGTCATTTAGTTCCAACTGCTACT 383

QY 428 TCATTTCCAGTGAAGAGAGGTTTGGTCTAAGAGTCAGCAGAACTCTGTGTGAGATGGAG 487
Db 384 TTATTTCTACTGAATCAGCATCTTGGCAAGACAGTCAGAAAGACTGTGCTAGAAATGGAG 443

QY 488 CACATTTGGTTGTTCAACACAGACAGCAGAGAAATTTCAATTTGCCAGCAGCTGAATG 547
Db 444 CTCACCTGCTGGTGATAACACTCAAGAGACAGAGGATTTCACTTCCAGAACTGCGAAG 503

QY 548 AGTCATTTTCTTATTTCTGGGGCTTTCAGACCCACAGGTAATAATAATTTGGCAATGGA 607
Db 504 AAGAATCTGCTTATTTGTGGGGCTCTCAGATCCAGAAGCTCAGCGACATTTGGCAATGG 563

QY 608 TTGATAGACACCTTATGAGAAATGTCAGATTTTGGCACTTGGCAGCTAGTGAGCCCAATCAT 667
Db 564 TTGATCAGACACCATACAATGAAAGTTCCACATTTCTGGCATCCACGTGAGCCAGTATC 623

QY 668 CTCAGACAGCATGCTTCAATAGTCTTCTGGAA---ACCTACAGGATGGGCTGGAATG 724
Db 624 CCAATGAGCGCTGCTGTGCTAAATTTTCGTAATCACCACAAAGATGGGCTGGAATG 683

QY 725 ATGTTATCTGTGAACACTAGAAGAAATTCATATGTGAGATGAATAAGATTTACCTATGAG 784
Db 684 ATGTTAATTGCTTGTGCTCCTCAAAGGTCAGTTTGTGAGATGATGAAGATCCACTTATGAA 743

QY 785 TAGAAGCTT 793
Db 744 CTGAACAT 752

RESULT 8
US-09-862-802-1
; Sequence 1, Application US/09862802
; Patent No. US20020165346A1
; GENERAL INFORMATION:
; APPLICANT: Schering-Plough Corporation
; TITLE OF INVENTION: MAMMALIAN MEMBRANE PROTEIN GENES; RELATED REAGENTS
; FILE REFERENCE: SF0695B
; CURRENT APPLICATION NUMBER: US/09/862,802
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/111,470
; PRIOR FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1104
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: mammalian nucleic acid
; NAME/KEY: CDS
; LOCATION: 242...952
; OTHER INFORMATION: protein coding sequence
US-09-862-802-1

Query Match      6.3%; Score 197; DB 9; Length 1104;
Best Local Similarity 67.8%; Pred. No. 7.2e-28;
Matches 291; Conservative 0; Mismatches 135; Indels 3; Gaps 1;

QY 368 AGGTGCCAGCCTGGGGATGTTGCCAGCTTCTTGGAAAGTCATTTGGTTCCAGTTGCTACT 427
Db 537 AAGAGACAGCCTGGAGCTGTTGCCAAAGAAATTGGAAGTCATTTAGTTCCAACTGCTACT 596
```





```

RESULT 12
US-09-833-381-1339
; Sequence 1339, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1339
; LENGTH: 756
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-381-1339

```

```

RESULT 13
US-09-965-529-40
; Sequence 40, Application US/09965529
; Publication No. US20020182671A1
; GENERAL INFORMATION:
; APPLICANT: LAL, Preeti
; APPLICANT: YUE, Henry
; APPLICANT: TANG, Y. Tom
; APPLICANT: BANDMAN, Olga
; APPLICANT: BURFORD, Neil
; APPLICANT: AZIMZAI, Yaida
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0731 USA
; CURRENT APPLICATION NUMBER: US/09/965,529
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/149,641; 60/164,203; PCT/US00/22315
; PRIOR FILING DATE: 1999-08-17; 1999-11-09; 2000-08-14
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PERL Program
; SEQ ID NO 40

```



Db 248 TCTTTCAAACCTGTGTGAGAAAAAGTT---TCAGCTACCTGAGAAATTTTCACAGAGCTCT 304  
Qy 347 CQTGCTTTCAGTGAAGGACAAAGGTGCCAGCTGGGGATGTGGCCAGCTTCTTTGGAAAGT 406  
Db 305 CQTGCTACAATTTATGATCA---GGTTCACTCAAGAAATGTTGTCATTTGAACCTGGGAAT 361  
Qy 407 CATTGGTTTCCAGTGTCTACTTTCATTTCCAGTGAAGAGAGGTTTGTCTTAAGAGTGAGC 466  
Db 362 ATTTTCAATCCAGCTGCTACTTCTTTTCTACTGACACCACTTTCTGTGGCGTTAAGTTTAA 421  
Qy 467 AGAAGTGTGTTGAGATGGGACACATTTGGTTGTGTTCAACACAGAACACAGCAGAAAT 526  
Db 422 AGAAGTGTGTCAGCCATGGGGGCTACCTGGTGGTTATCAACTCAAGAGGAGCAGGAAT 481  
Qy 527 TCATTGTCCAGCAGCTGAATGAGTCAATTTCTTATTTCTGGGGCTTTCAGACCCACACAAG 586  
Db 482 TCCTTTCTTCAAGAAACCTAAATGACAGAGCTTTTATTTGGAGTGTCAACACAGGTTG 541  
Qy 587 GTAATAATTAATTTGGCAATGGATTGATAAGACACACCTTATGAGAAAAATGTCAGATTTTGGC 646  
Db 542 TCGAGGGTCAGTGGCAATGGGTGGACGGCACACCTTTTGACAAAGTCTCTGAGCTTCTGGG 601  
Qy 647 ACCTAGGTGAGCCCAATCA-----TTCTGCAGACAATGTGCTTCAATAGTCTTCTGGA 700  
Db 602 ATGTAGGGGAGCCCAACACATAGTACCTGGAGGAGCTGTGCCACCATGAGAGACTCTT 661  
Qy 701 AAACCTACAGGATGGGCTGGAATGATGTTATCTGTGAAACTAGAAGGAATTCAAATATG 760  
Db 662 CAAACCAAGGCAAAATTTGGAATGATGTAACCTGTTTCTCTCAATTTTTCGGATTG 721  
Qy 761 AGATG 765  
Db 722 AAATG 726

RESULT 15  
US-09-902-853-376  
; Sequence 376, Application US/09902853  
; Publication No. US20020192659A1  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc.  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, A.  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth, J.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Mather, Jennie P.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William, I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: 10466-14  
; CURRENT APPLICATION NUMBER: US/09/902,853  
; CURRENT FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: US/09/665,350  
; PRIOR FILING DATE: 2000-09-18

; PRIOR APPLICATION NUMBER: US 60/143,048  
; PRIOR FILING DATE: 1999-07-07  
; PRIOR APPLICATION NUMBER: US 60/145,698  
; PRIOR FILING DATE: 1999-07-26  
; PRIOR APPLICATION NUMBER: US 60/146,222  
; PRIOR FILING DATE: 1999-07-28  
; PRIOR APPLICATION NUMBER: PCT/US99/20594  
; PRIOR FILING DATE: 1999-09-08  
; PRIOR APPLICATION NUMBER: PCT/US99/20944  
; PRIOR FILING DATE: 1999-09-13  
; PRIOR APPLICATION NUMBER: PCT/US99/21090  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR APPLICATION NUMBER: PCT/US99/21547  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR APPLICATION NUMBER: PCT/US99/23089  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR APPLICATION NUMBER: PCT/US99/28214  
; PRIOR FILING DATE: 1999-11-29  
; PRIOR APPLICATION NUMBER: PCT/US99/28313  
; PRIOR FILING DATE: 1999-11-30  
; PRIOR APPLICATION NUMBER: PCT/US99/28564  
; PRIOR FILING DATE: 1999-12-02  
; PRIOR APPLICATION NUMBER: PCT/US99/28565  
; PRIOR FILING DATE: 1999-12-02  
; PRIOR APPLICATION NUMBER: PCT/US99/30095  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: PCT/US99/30911  
; PRIOR FILING DATE: 1999-12-20  
; PRIOR APPLICATION NUMBER: PCT/US99/30999  
; PRIOR FILING DATE: 1999-12-20  
; PRIOR APPLICATION NUMBER: PCT/US00/00219  
; PRIOR FILING DATE: 2000-01-05  
; NUMBER OF SEQ ID NOS: 423  
; SEQ ID NO 376  
; LENGTH: 997  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-09-902-853-376

Query Match 3.8%; Score 117.4; DB 9; Length 997;  
Best Local Similarity 54.4%; Pred. No. 4.2e-13;  
Matches 329; Conservative 0; Mismatches 261; Indels 15; Gaps 4;  
Qy 170 AACCTCAAAGTACAGAAAGAGGGTGTGTCCTCC---TGAGACTCTGTGTGGCTG 226  
Db 128 AACACAATGCACAGAGAGAGGATCTCTCTCCCAAAATGTTCTTATGAGACTGTGCTG 187  
Qy 227 GGATTTCCATTCACCTCCTCAGTGTGCTTCATTTGTGAGCTGTGTAGTAACCTACCAAT 286  
Db 188 GGATCCCCCATCTTCTTCAGTGCCCTGTTTCATCACCCAGATGTGTGTGACATTTCCGCA 247  
Qy 287 TTACATATGGTGAACCTGGCAAAAGGCTGTCTGAACCTACACTCATATCATTCATTAAGTCTCA 346  
Db 248 TCCTTCAAACCTGTGATGAGAAAAAGTT---TCAGCTACCTGAGAAATTCACAGAGCTCT 304  
Qy 347 CQTGCTTTCAGTGAAGGACAAAGGTGCCAGCTGGGGATGTTGCCAGCTTCTTTGGAAAGT 406  
Db 305 CQTGCTACAATTTATGATCA---GGTTCACTCAAGAAATTTGTTCCATTTGAACCTGGGAAT 361  
Qy 407 CATTGGTTCAGTGTGCTACTTTCATTTCCAGTGAAGAGAGGTTTGTGCTTAAGAGTGAGC 466  
Db 362 ATTTTCAATCCAGTGTGCTACTTCTTTCTACTGACACCACTTCTCTGGCGTTAAGTTTAA 421  
Qy 467 AGAAGTGTGTTGAGATGGGAGCACATTTGGTTGTGTTCAACACAGAACACAGCAGAAAT 526  
Db 422 AGAAGTGTGAGATGGGAGCACATTTGGTTGTGTTCAACACAGAACACAGCAGAAAT 481  
Qy 527 TCATTTGCCAGAGCTGAATGAGTCAATTTCTTATTTTCTGGGGCTTTCAGACCCCAAG 586  
Db 482 TCCTTTCTCTACAAGAAACCTTAAATGAGAGAGTGTGTTTATTTGAGACTCTCTCAGACAGTTG 541  
Qy 587 GTATAATAATTTGGCAATGCAATTCATAAGACACCTTATGAGAAAAATGTCAGATTTTGGC 646





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 17, 2003, 08:47:16 : Search time 2991.64 Seconds  
(without alignments)  
16857.906 Million cell updates/sec

Title: US-09-766-511B-51  
Perfect score: 3114  
Sequence: 1 cttaatgttggaagtctctt.....tttaaaaaaaaaaaaaaa 3114

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

1:	em_estba:*
2:	em_esthum:*
3:	em_estin:*
4:	em_estnu:*
5:	em_estov:*
6:	em_estpl:*
7:	em_estro:*
8:	em_htc:*
9:	gb_est1:*
10:	gb_est2:*
11:	gb_htc:*
12:	gb_est3:*
13:	gb_est4:*
14:	gb_est5:*
15:	em_estfun:*
16:	em_estom:*
17:	gb_gss:*
18:	em_gss_hum:*
19:	em_gss_inv:*
20:	em_gss_pln:*
21:	em_gss_vrt:*
22:	em_gss_fun:*
23:	em_gss_mam:*
24:	em_gss_mus:*
25:	em_gss_other:*
26:	em_gss_pro:*
27:	em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	366.8	11.8	383	17	AQ091895 HS_3014.B
2	366.8	11.8	1900	11	BC001229 Homo sapi
3	366.2	11.8	563	10	AW576568 UI-HF-BN0
4	365.8	11.7	582	13	BM505485 ih12e08.x
5	364.6	11.7	732	14	BM972330 UI-CF-EC1
6	364.6	11.7	764	14	BQ009055 UI-H-E10-

7	364.2	11.7	511	9	A1745635
8	363	11.7	1030	13	BM473183
c 9	361.4	11.6	588	13	BT868214
c 10	360.6	11.6	1079	13	BM463804
11	359.8	11.6	538	10	AW576183
12	358.2	11.5	548	9	A1684386
13	358.2	11.5	577	14	BM968219
14	355.6	11.4	725	14	BM715473
15	355	11.4	496	9	AA557129
16	354	11.4	528	13	BM661657
17	354	11.4	735	14	BM678541
18	354	11.4	736	14	BM679660
19	354	11.4	762	14	BM679029
20	352.2	11.3	519	14	C75436
21	350.2	11.2	496	9	A1025096
22	350.2	11.2	510	9	A1018792
23	349.8	11.2	596	9	AA172069
24	349.8	11.2	599	9	AA172043
25	349.6	11.2	488	10	BE350097
26	349.2	11.2	506	10	AW148744
27	349	11.2	487	10	BE302211
28	347.2	11.1	578	9	AU149516
29	346.8	11.1	600	9	AU145485
30	346.4	11.1	652	9	AA172077
c 31	344.6	11.1	534	14	BM754950
32	344.2	11.1	590	9	AU157092
33	343.6	11.0	504	12	BF603486
34	343.6	11.0	543	9	AU152484
35	342.6	11.0	477	9	A1924217
36	341.8	11.0	512	9	A1422372
c 37	341.6	11.0	480	14	BM832889
38	340.6	10.9	545	10	BE177944
39	338.8	10.9	553	13	BM145299
40	338.8	10.9	573	9	AU151353
41	338.6	10.9	477	9	A1040207
42	338.6	10.9	477	9	A1660533
43	338	10.9	545	13	BG942747
44	337.8	10.8	535	9	AU153279
45	335.4	10.8	465	9	A1246452

ALIGNMENTS

RESULT 1	AQ091895	383 bp	DNA	linear	GSS 27-AUG-1998
AQ091895	HS_3014_B2_C05_T7 CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3014 Col=10 Row=F, DNA sequence.				
LOCUS	AQ091895				
DEFINITION	AQ091895				
ACCESSION	AQ091895.1				GI:3462982
VERSION	GSS.				
KEYWORDS	human.				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 383)				
AUTHORS	Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.				
TITLE	Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)				
MEDLINE	99380589				
COMMENT	Contact: Mahairas GG, Wallace JC, Hood L High Throughput Sequencing Center University of Washington 401 Queen Anne Avenue North, Seattle, WA 98109, USA Tel: (206) 616-3618 Fax: (206) 616-3887 Email: jwallace@u.washington.edu Sequence Tagged Connector Plate: 3014 row: F column: 10				

Class: BAC ends  
High quality sequence stop: 383.  
Location/Qualifiers  
1. 383  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="Plate-3014 Col-10 Row-F"  
/clone\_lib="CIT Approved Human Genomic Sperm Library D"  
/sex="male"  
/note="Organ: sperm; Vector: pBelOBAC11; BAC Clones in E-Coli DH10B"  
BASE COUNT 110 a 72 c 49 g 148 t 4 others  
ORIGIN

Query Match 11.8%; Score 366.8; DB 17; Length 383;  
Best Local Similarity 97.1%; Pred. No. 1.4e-46;  
Matches 371; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
QY 843 TTAATAATGAACCCATCATGAATGATAATTTCTTCTGATTTTACACATAATCCCTTAT 902  
Db 1 TTAANATGAACCCATCATGAATGATAATTTCTTCTGATTTTACACATAATCCCTTAT 60

QY 903 GTTATAGAGTTTCACAGAAATGAAGATACCTGTTTCCCTTTAATCAATCTTCTCGTTT 962  
Db 61 GTTATAGAGTTTCACAGAAATGAAGATACCTGTTTCCCTTTAATCAATCTTCTCGTTT 120  
QY 963 CCTCTTTTTCATTAATGATAGAAATGCACCCCTTCTCTTTGTTCCATTTCTTCACTTGT 1022  
Db 121 CCTCTTTTTCATTAATGATAGAAATGCACCCCTTCTCTTTGTTCCATTTCTTCACTTGG 180

QY 1023 TATTCATTTTTTCTTCTTCCACTTCATTTACACAAATATTTATGTTTTCAGAGACGT 1082  
Db 181 TATTCATTTTTTCTTCTTCCACTTCATTTACACAAATATTTATGTTTTCAGAGACGT 240  
QY 1083 ACTATTTGTTTGTAGAGATTTATAAGGAGATCTTTTGCAAAATATGACTTTCCTT 1142  
Db 241 NACTATTTTGTGTTAGAGATTTATAAGGAGATCTTTTGCAAAATATGACTTTCCTT 300

QY 1143 CCTCAATATACCATAAAGAAATCTTTTGGTCAAGATGGTAGTGGAACTACAATCATCT 1202  
Db 301 CCTCAATATACCATAAAGAAATCTTTTGGTCAAGATGGTAGTGGAACTACAATCATCT 360  
QY 1203 GAAGGCTGACAAGAGTTGAAA 1224  
Db 361 GAAGGCTGACAAGAGTTGAAA 382

RESULT 2  
BC001229/c 1900 bp mRNA linear HTC 31-JAN-2002  
LOCUS Homo sapiens, Similar to RAD21 (S. pombe) homolog, clone  
DEFINITION IMAGE:3455406, mRNA.

ACCESSION BC001229  
VERSION BC001229.1 GI:12654772  
KEYWORDS HTC.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1900)

REFERENCE Strausberg, R.  
AUTHORS Direct Submission  
TITLE Submitted (11-DEC-2000) National Institutes of Health, Mammalian  
JOURNAL Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA

REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov  
COMMENT Contact: MGC help desk  
Email: cgapbs-re@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome

Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: http://www.shgc.stanford.edu  
Contact: (Dickson, Mark) mcd@paxil.stanford.edu  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Series: IRAK Plate: 4 Row: 1 Column: 9  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7705806  
This clone has the following problem: no 5' EST match.

FEATURES  
Location/Qualifiers  
1. 1900  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3455406"  
/tissue\_type="Cervix, carcinoma"  
/clone\_lib="NIH\_MGC\_12"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
BASE COUNT 632 a 294 c 362 g 612 t  
ORIGIN

Query Match 11.8%; Score 366.8; DB 11; Length 1900;  
Best Local Similarity 86.4%; Pred. No. 5.8e-47;  
Matches 472; Conservative 0; Mismatches 47; Indels 27; Gaps 5;  
QY 2477 TTTCCTTTCTCTGCTCCAAAATATTTTTCACAGCTTATAACACAACTTTTATTAGA 2536  
Db 1896 TTTTCTTTTCTTTTCTTTTCTTTTTCACAGCTTATAACACAACTTTTATTAGA 1837

QY 2537 AAAGTTATACATAACAGCATCAACTATTTTCAAG-----AACCCAATAGCAA 2586  
Db 1836 AAAGTTATACATAACAGCATCAACTATTTTCAAGACAATATTAAACCCGATAGCAA 1777

QY 2587 CAAAAACAGACTAACAAAATGTTAACAGAACTTAATGACCTTTCTTAAATCAAAAT 2646  
Db 1776 CAAAAACAGACTAACAAAATGTTAACAGAACTTAATGACCTTTCTTAAATCAAAAT 1717

QY 2647 TCAATTATCTACAAATGCTATTTTACAAACAGGAAACTCCATGTTTACAGGCATGCA 2706  
Db 1716 TCAATTATCTACAAATGCTATTTTACAAACAGGAAACTCCATGTTTACAGGCATGCA 1657

QY 2707 TATTGAAATAAAGCTGCAATAGC-TTTTATACAAATTCGCTCTCAAGAAATGAATC 2765  
Db 1656 TATTGAAATAAAGCTGCAATAGC-TTTTATACAAATTCGCTCTCAAGAAATGAATC 1597

QY 2766 ATTAAGACAGTAATAGGAGTTTACAAAATTTAAACATTTTCAAGTAATTTTAAATATTG 2825  
Db 1596 ATTAAGACAGTAATAGGAGTTTACAAAATTTAAACATTTTCAAGTAATTTTAAATATTG 1537

QY 2826 TCTTCAATAATTTTAAATTTTCAAGCTGAGTTTCAAAAGTGA---TTTTTCCCAAA 2882  
Db 1536 GGT-----ATACACTGAAGCTGAGTTTCAAAAGTGAATTTTTCACCAAA 1489

QY 2883 AGGTGCCAACACTTAAGCTAGAGCTTTTCAAGTGTAACTTTTGCCT-AAAAGTTTAAGACAT 2941  
Db 1488 AAGTTTCAACACTTAAGCTAGAGCTTTTCAAGTGTAACTTTTGCCTAAAAGTTTAAGACAT 1429

QY 2942 ATTCTGAGAATCATATAGTCACATGATTTCTGATGCTATCTGCTCTGTTAATAACAAG 3001  
Db 1428 ATTCTGAGAATCATATAGTCACATGATTTCTGATGCTATCTGCTCTGTTAATAACAAG 1369  
QY 3002 ATTTC A 3007  
Db 1368 TCITTA 1363

RESULT 3  
AW576568 563 bp mRNA linear EST 15-MAR-2000  
LOCUS AW576568  
DEFINITION UI-HF-BN0-alu-a-09-0-UI-sl NIH\_MGC\_50 Homo sapiens cDNA clone



IMAGE:3080729 3', mRNA sequence.  
ACCESSION AW576568  
VERSION AW576568.1 GI:7248107  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NIH-MGC <http://mgc.nci.nih.gov/>.  
1 (bases 1 to 563)  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: rcapbs@remail.nih.gov  
The sequence contained an oligo-dT track that was present in the  
oligonucleotide that was used to prime the synthesis of first  
strand cDNA and therefore this may represent a bonafide poly A  
tail. Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
cDNA Library Preparation: M.B. Soares Lab  
cDNA Library Arrayed by: M.B. Soares Lab  
DNA Sequencing by: M.B. Soares Lab  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/URL at:  
[www-bio.lnl.gov/Dbp/image/image.html](http://www-bio.lnl.gov/Dbp/image/image.html)  
Seq primer: M13 Forward  
POLYA=Yes.  
FEATURES  
source  
Location/Qualifiers  
1..563  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3080729"  
/clone\_lib="NIH\_MGC\_50"  
/tissue\_type="lymph"  
/cell\_type="germinal center B cells"  
/cell\_line="MGC85"  
/lab\_host="DH10B (LTI)"  
/note="Vector: pT73-Pac; Site.1: NotI; Site.2: Eco RI;  
Constructed from size fractionated cytoplasmic mRNA  
(3.5-4.4Kb). Directionally cloned. Cells provided by  
Louis M. Staudt, Ph.D. Library preparation by Maria de  
Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."  
BASE COUNT 215 a 91 c 63 g 194 t  
ORIGIN  
Query Match 11.8%; Score 366.2; DB 10; Length 563;  
Best Local Similarity 88.5%; Pred. No. 1.4e-46;  
Matches 463; Conservative 0; Mismatches 33; Indels 27; Gaps 5;  
QY 2500 TATTTTTTTTACAGCTTATACACAACTTTTATTAGAAAAGTTATACATAACACAGCATC 2559  
DB 8 TTTTTTTTTTACAGCTTATACACAACTTTTATTAGAAAAGTTATACATAACATAGCATC 67  
QY 2560 AACTATTTTTCAG-----AACCCAATAAGCAACAAAAACACAGACTAACAAAATGT 2609  
DB 68 AACTATTTTTCAGAAACAATAATTAAACCCGATAAGCAACAAAAACACAGACTAACAAAATGT 127  
QY 2610 GTACAGAAACTAATGACCTTCTTAAATCAACATTCATTTATCTACAACTCTATT 2669  
DB 128 GTACAGAAACTAATGACCTTCTTAAATCAACATTCATTTATCTACAACTCTATT 187  
QY 2670 ACAACAGGGAACCTCCATGGTTTTCACAGGCATGTCATATTGAAAATAAGCTGCAATAG 2729  
DB 188 ACAACAGGGAACCTCCTGGTTTTCACAGGCATCATATTGAAATATAAGCTGCAATAG 247  
QY 2730 C-TTTTATACAATTATCGCTCTCAAGAAAATGAATCATTAAGACAGTAAATTAGGATTC 2788  
DB 248 CAATTTTATACAATTACCACCTCTGAAGAAACTGAATCATTTAAACACAGTAATTACGATTC 307  
QY 2789 ACAATTTTAAACATTTTCAGCTAAATTTTAAATTTATGCTTCTTAATTTTAAATTTATG 2848  
DB 308 ACAATTTTAAACATTTTTCACATAAATTTTAAATTTATTTGGGT-----ATACACTG 355  
QY 2849 AAGTCTGAGTTTCAAAAGTGA---TTTTTTCCCAAAAGGTGCCAACACTTAAGCTTAGAG 2905

|||||  
Db 356 AAGTCTGAGTTTCAAAAGTGATTTTTTTTCCCAAAAGTTTCAACACTTAAGCTACAA 415  
QY 2906 CTTTCAGTGTAACTTTGCCCT-AAAGTTAAGACATATTTCTGAGAATCATATAAGTCAC 2964  
Db 416 CTTTCAGTGTAACTTTGCCCTTAAAGTTAAGACATATTTCTGATAATCATATAACAGTCAC 475  
QY 2965 ATGATTTCTGATGCTATCTGCTCTCTTAATAACAAGATTTC 3007  
Db 476 ATGATTTCTGATGCTATCTGCTCTCTTAATAACAAGCTTTTA 518  
RESULT 4  
BM505485  
LOCUS BM505485 582 bp mRNA linear EST 14-FEB-2002  
DEFINITION ih12e08.x1 Human insulinoma Homo sapiens cDNA 3', mRNA sequence.  
ACCESSION BM505485  
VERSION BM505485.1 GI:18668752  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 582)  
Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,  
Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,  
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,  
Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas  
M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R., Williams, T.,  
Jackson, Y. and Bowers, Y.  
Endocrine Pancreas Consortium  
Unpublished (2000)  
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
MA 02138  
Tel: 617-495-1812  
Fax: 617-495-8557  
Email: dmelton@bioh.harvard.edu  
Library was constructed by Dr. J. H. Inoue DNA sequencing by: Washington  
pBluescript SK- by Dr. H. Inoue  
University Genome Sequencing Center For information on obtaining a  
clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)  
Seq primer: 40UP from Gibco  
High quality sequence stop: 473.  
Location/Qualifiers  
1..582  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="Human insulinoma"  
/tissue\_type="insulinoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: pancreas; Vector: pBluescript SK-; Site.1:  
XhoI; Site.2: EcoRI; Constructed with lambda ZAPII system  
(Stratagene) by Dr. J. H. Inoue following the Washington  
pBluescript SK- by Dr. H. Inoue protocol  
University protocol  
([http://genome.wustl.edu/est/lambda\\_protocol.shtml](http://genome.wustl.edu/est/lambda_protocol.shtml)).  
Please contact Hiroshi Inoue, MD/PhD for further  
information on this library (Metabolism Division, Permutt  
Laboratory, Washington University School of Medicine, Box  
8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this  
is a Washington University Pancreas EST project library."  
BASE COUNT 221 a 95 c 72 g 194 t  
ORIGIN  
Query Match 11.7%; Score 365.8; DB 13; Length 582;  
Best Local Similarity 88.7%; Pred. NO. 1.6e-46;  
Matches 462; Conservative 0; Mismatches 32; Indels 27; Gaps 5;  
QY 2502 TTTTTTTTACAGCTTATACACAACTTTTATTAGAAAAGTTATACATAACACAGCATCAA 2561  
|||||



```

VERSION      BQ009055.1  GI:19733956
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE        1 (bases 1 to 764)
JOURNAL      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
COMMENT      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
              Tumor Gene Index
              Unpublished (1997)
              Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-r@mail.nih.gov
              Tissue Procurement: Dr. Jose Mercuende
              cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
              DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
              Clone Distribution: Clone distribution information can be found
              through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
              Seq primer: M3 FORWARD
              POLYA=yes.

FEATURES     Location/Qualifiers
             1..764
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /clone_lib="IMAGE:5840869"
             /clone_lib="NCI-CGAP_E10"
             /tissue_type="Chondrosarcoma"
             /dev_stage="Adult"
             /lab_host="DH10B (Life Technologies)"
             /note="Organ: Left Pelvis; Vector: pT7T3-Pac (Pharmacia)
             with a modified polylinker; Site.1: EcoR I; Site.2: Not I;
             NCI-CGAP_E10 is a cDNA library containing the following
             tissue(s): Chondrosarcoma. The library was constructed
             according to Bonaldo, Lennon and Soares, Genome Research,
             6:791-806, 1996. First strand cDNA synthesis was primed
             with an oligo-dT primer containing a Not I site. Double
             stranded cDNA was ligated to an EcoR I adaptor, digested
             with Not I, and cloned directionally into pT7T3-Pac
             vector. The oligonucleotide used to prime the synthesis of
             first-strand cDNA contains a library tag sequence that is
             located between the Not I site and the (dT)18 tail. The
             sequence tag for this library is ACATTGTCAC.
             TAG_LTB=UI-H-E10
             TAG_TISSUE=chondrosarcoma
             TAG_SEQ=ACATTGTCAC"

BASE COUNT   272 a 131 c 107 g 253 t 1 others
ORIGIN
Query Match      11.7%; Score 364.6; DB 14; Length 764;
Best Local Similarity 88.3%; Pred. No. 2e-46;
Matches 462; Conservative 0; Mismatches 34; Indels 27; Gaps 5;

Qy 2500 TATTTTTTTTACAGCTTATACACAACTTTTATTAGAAAAGTTATACATACACAGCATC 2559
Db 9 TTTTTTTTTTACAGCTTATACACAACTTTTATTAGAAAAGTTATACATACATAGCATC 68

Qy 2560 AACTATTTTTCAG-----AACCAATAAGCACAAAAACCCAGACTAACAAAATGT 2609
Db 69 AACTATTTTTCAGAACAAATATTAACCCGATAGACAAAAAACCCAGACTAACAAAATGT 128

Qy 2610 GTAACAGAAACTAATGACCTTCTTAAATCAACATTCATTTATCTACATATGCTATTT 2669
Db 129 GTACAGAAACTAATGACCTTCTTAAATCAACATTCATTTATCTACATGCTTTT 188

Qy 2670 ACAACAGGGGAAACTCCATGGTTTACAGGCATGTCATATTTGAAATTAAGCTGCAATAG 2729
Db 189 ACAACAGGGGAAACTCCTGGTTTACAGGCATCATATTTGAATATATAAGCTGCAATAG 248

Qy 2730 C-TTTTATACAAATATCGCTCTCAAGAAATGAATCATTAAGACAGTAATTAGGAGTTC 2788
Db 249 CAATTTTATACAAATTAGCACTCTGAAGAACTGAATCATTTAAACACAGTAATTAGGAGTTC 308

Qy 2789 ACAAATTTAAACATTTTCACGTAATTTTAAATTTATCTTCAATAAATTTTAAATTTT 2848
Db 309 ACAAATTTAAACATTTTCACATAATTTTAAATTTATTTGGGT-----ATACACTG 356

Qy 2849 AAGTCTGAGTTTCAAAAAGTGA---TTTTTCCCAAGAGTGCCCAACACTTAAAGCTAGAG 2905
Db 357 AAGTCTGAGTTTCAAAAAGTGAATTTTTCCTCCCAAGAGTTTCAACACTTAAAGCTAGAA 416

Qy 2906 CTTTCAGTGTAACTTTGCCCC-AAAAGTTAAGACATATTTCTGAGATCATATAATAGTCAC 2964
Db 417 CTTTCAGTGTAACTTTGCCCC-AAAAGTTAAGACATATTTCTGATATCATACAGTCAC 476

Qy 2965 ATGATTTCTGATGCTATCTGCTGTTTAAATAACAAAGATTTC 3007
Db 477 ATGATTTCTGATGCTATCTGCTGTTTAAATAAAGTCTTTA 519

RESULT 7
LOCUS     AI745635
DEFINITION tr24a09.xl NCI-CGAP_Ov23 Homo sapiens cDNA clone IMAGE:2219224 3',
           mRNA sequence.
ACCESSION AI745635
VERSION    AI745635.1 GI:5113923
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL    National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT    Tumor Gene Index
           Unpublished (1997)
           Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-r@mail.nih.gov
           Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
           Emmert-Buck, M.D., Ph.D.
           cDNA Library Preparation: Life Technologies, Inc.
           DNA Sequencing by: Greg Lennon, Ph.D.
           Clone Distribution: NCI-CGAP clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           www-bio.llnl.gov/bbrp/image/image.html
           Seq primer: -400P from Gibco
           High quality sequence stop: 414.
           Location/Qualifiers
           1..511
           /organism="Homo sapiens"
           /db_xref="taxon:9606"
           /clone_lib="IMAGE:2219224"
           /clone_lib="NCI-CGAP_Ov23"
           /tissue_type="tumor, 5 pooled (see description)"
           /lab_host="DH10B"
           /note="Organ: ovary; Vector: pCMV-SPORT6; Site.1: SalI;
           Site.2: NotI; Cloned unidirectionally. Primer: Oligo dT.
           Average insert size 1.35 kb. Tumor types include: mixed
           Mullerian tumor, papillary serous, clear cell, spindle
           cell. All are primary tumors, metastasis positive. Life
           Technologies catalog #: 11534-013"

BASE COUNT   199 a 87 c 57 g 168 t
ORIGIN
Query Match      11.7%; Score 364.2; DB 9; Length 511;
Best Local Similarity 88.5%; Pred. No. 2.9e-46;
Matches 461; Conservative 0; Mismatches 33; Indels 27; Gaps 5;

Qy 2502 TTTTTTTTACAGCTTATAACACAACTTTTATTAGAAAAGTTATACATACACAGCATCAA 2561
Db 1 TTTTTTTTACAGCTTATAACACAACTTTTATTAGAAAAGTTATACATACATAGCATCAA 60

Qy 2562 CTATTTTTCAG-----AACCAATAAGCACAAAAACCCAGACTAACAAAATGTCT 2611
Db 1 TTTTTTTTTCAG-----AACCAATAAGCACAAAAACCCAGACTAACAAAATGTCT 2611

```

Db	61	CTATTTTCAAGACAATATTAAACCCGATAGCMACAAAACAGACTAACAAAATGTGT	120
Qy	2612	AACAGAAACTAATGACCTTTCTAAATCAAACTTCAATATCTACAATGTCTATTAC	2671
Db	121	AACAAGAACTAATGACCTTTCTAAATCAAACTTCAATATCTACAATGTCTATTAC	180
Qy	2672	AAACAGGAAACCTCCATGTTTACAGGCATGTCATATTGAAATAAAGCTGCAATAGC-	2730
Db	181	AAACGGGAAACCTCCATGTTTACAGGCATGTCATATTGAAATAAAGCTGCAATAGCA	240
Qy	2731	TTTTTATACAAATATCGCTCTCAAGAAATGAATCATTAAGACAGTAATTAGAGTTCAC	2790
Db	241	ATTTTATACAAATATCGCTCTCAAGAAATGAATCATTAAGACAGTAATTAGAGTTCAC	300
Qy	2791	AAATTTAAACATTTACAGTAAATTTTAAATATTGTTCTTCAATAATTTTAAATTAATTGAA	3850
Db	301	AAATTTAAACATTTACATAAATTTTAAATATTGTTCTTCAATAATTTTAAATTAATTGAA	348
Qy	2851	GTCTGAGTTTCAAAAGTCA--TTTTTCCCAAAAGGTCGCAACACTTAAGCTAGAGCT	2907
Db	349	GTCTGAGTTTCAAAAGTCA--TTTTTCCCAAAAGGTCGCAACACTTAAGCTAGAGCT	408
Qy	2908	TTCACTGTTTAACTTTGCGCT-AAAAGTTAAGACATATTTCTGAGAATCATAATAGTCACAT	2966
Db	409	TTCACTGTTTAACTTTGCGCT-AAAAGTTAAGACATATTTCTGAGAATCATAATAGTCACAT	468
Qy	2967	GATTTCTGATGCTATCTGCTCTGTTAATAACAAGATTTCAC	3007
Db	469	GATTTCTGATGCTATCTGCTCTGTTAATAACAAGATTTCAC	509
RESULT 8			
LOCUS	BM473183	1030 bp	mRNA
DEFINITION	AGENCOURT_6462477 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5519713	5', mRNA sequence.	
ACCESSION	BM473183		
VERSION	BM473183.1	GI:18522225	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 1030)		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabbs-remail.nih.gov Tissue procurement: ATCC CDNA Library Preparation: Life Technologies, Inc. DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM12182 row: j column: 02 High quality sequence stop: 604.		
FEATURES	Location/Qualifiers		
source	1..1030		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5519713"		
	/clone_lib="NIH_MGC_71"		
	/tissue_type="leiomyosarcoma"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: Not I; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2.1 kb."		
BASE COUNT	365 a 215 c 135 g 315 t		
ORIGIN			
Query Match	11.7%;	Score 363;	DB 13; Length 1030;

Best Local Similarity 88.1%; Pred. No. 3.le-46;			
Matches	461;	Conservative	0; Mismatches 35; Indels 27; Gaps 5;
Qy	2500	TATTTTTCACAGCTTATACACAACTTTTATTAGAAAAGTTATACATACACAGCATC	2559
Db	9	TTTTTTTTTACCGCTTATACACAACTTTTATTAGAAAAGTTATACATACATAGCATC	68
Qy	2560	AACTATTTTCAAG-----AACCCAATAAGCAACAAAACAGACTAACAAAATGT	2609
Db	69	AACTATTTTCAAGAACTATTTAAACCCGTAAGCAACAAAACAGACTAACAAAATGT	128
Qy	2610	GTACAAGAACTAATGACCTTTCTAAATCAAACTTCAATATCTACAATGTCTATT	2669
Db	129	GTACAAGAACTAATGACCTTTCTAAATCAAACTTCAATATCTACAATGTCTATT	188
Qy	2670	ACAAACAGGAAACCTCCATGTTTACAGGCATGTCATATTGAAAATAAAGCTGCAATAG	2729
Db	189	ACAAACGGGAAACCTCCATGTTTACAGGCATGTCATATTGAAAATAAAGCTGCAATAG	248
Qy	2730	C-TTTTATACAAATATCGCTCTCAAGAAAATGAATCATTAAGACAGTAATTAGAGTTC	2788
Db	249	CAATTTTATACAAATATCGCTCTCAAGAAAATGAATCATTAAGACAGTAATTAGAGTTC	308
Qy	2789	ACAAATTTAAACATTTCACTGTAATTTTAAATATTGTTCTTCAATAATTTTAAATATTG	2848
Db	309	ACAAATTTAAACATTTCACTGTAATTTTAAATATTGTTCTTCAATAATTTTAAATATTG	356
Qy	2849	AAGCTGAGTTTCAAAAGTGA--TTTTTCCCAAAAGGTCGCAACACTTAAGCTAGAG	2905
Db	357	AAGCTGAGTTTCAAAAGTGA--TTTTTCCCAAAAGGTCGCAACACTTAAGCTAGAG	416
Qy	2906	CTTTCAGTGTAACTTTGCCCT-AAAAGTTAAGACATATTTCTGAGAATCATAATAGTCAC	2964
Db	417	CTTTCAGTGTAACTTTGCCCT-AAAAGTTAAGACATATTTCTGAGAATCATAATAGTCAC	476
Qy	2965	ATGATTTCTGATGCTATCTGCTCTGTTAATAACAAGATTTCAC	3007
Db	477	ATGATTTCTGATGCTATCTGCTCTGTTAATAACAAGATTTCAC	519
RESULT 9			
LOCUS	BI868214	588 bp	mRNA
DEFINITION	603392157F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5402031 5',		
ACCESSION	BI868214		
VERSION	BI868214.1	GI:16041887	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 588)		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabbs-remail.nih.gov Tissue procurement: ATCC CDNA Library Preparation: Life Technologies, Inc. DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL) Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM12025 row: b column: 16 High quality sequence stop: 588.		
FEATURES	Location/Qualifiers		
source	1..588		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5402031"		
	/clone_lib="NIH_MGC_90"		



**AUTHORS** NIH-MGC <http://mgc.nci.nih.gov/>.  
**TITLE** National Institutes of Health, Mammalian Gene Collection (MGC)  
**JOURNAL** Unpublished (1999)  
**COMMENT** Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. Tissue procurement: Louis M. Staudt, M.D., Ph.D.  
cDNA Library Preparation: M.B. Soares Lab  
cDNA Library Arrayed by: M.B. Soares Lab  
DNA Sequencing by: M.B. Soares Lab  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: [www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
Seq primer: M13 Forward  
POLYA=Yes.

FEATURES	source	location/Qualifiers
1..538		/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/clone="IMAGE:3065073"
		/clone_lib="NIH_MGC_50"
		/tissue_type="lymph"
		/cell_type="germinal center B cells"
		/cell_line="MGC85"
		/lab_host="DH10B (LTI)"
		/note="vector: pT73-Pac; Site_1: NotI; Site_2: Eco RI; Constructed from size fractionated cytoplasmic mRNA (3.5-4.4kb). Directionally cloned. Cells provided by Louis M. Staudt, Ph.D. Library preparation by Maria de Fatima Bonaldo, ph.D. and M. Bento Soares, ph.D."
BASE COUNT	201 a	91 c 61 g 185 t
ORIGIN		
	Query Match	11.6%; Score 359.8; DB 10; Length 538;
	Best Local Similarity	87.8%; Pred. No. 1.3e-45;
	Matches 459; Conservative	0; Mismatches 37; Indels 27; Gaps 5;
QY	2500	TATTTTTTTACAGCTTATAACACAACCTTTTATTAGAAAAGTTATACATTAACACAGCATC 2559
Db	8	
		TTTTTTTTTACAGCTTATAACACAACCTTTTATTAGAAAAGTTATACACAAATACATC 67
QY	2560	AACATATTTTCAAG-----AACCCAAATAAGCAACAAAAACCAGACTTAACAAAATGT 2609
Db	68	AACTATTTTCAAGAACAATATTAACCGGATAGCAACAAHAACCGACTTACAAATGT 127
QY	2610	GTAACAAGAACTAATGACCTTCTAAAYCAAACATTCATTAATTCACAATGTCTATTTT 2669
Db	128	GTAAAGAAGAACTAATGACCTTCTTAAATCAAAACATTCATTAATTCACAATGTCTTTT 187
QY	2670	ACAAACAGGGAAACCTCATGGTTTACAGGCACTGCATATTGAAATTAAGCTGCAATAG 2729
Db	188	ACAAACGGGAACACTCTTGTTTACAGGCACATCATTTGAATATTAAGCTGCAATAG 247
QY	2730	C-TTTTATFACAATTTATCGCTCTCAAGAAATGAATCATTTAAGACAGTAATTAGGAGTTC 2788
Db	248	CAATTTTATACAAATTTACCACCTCTGAAGAACTGAATCATTTAAAACACAGTAATACGAGTTC 307
QY	2789	ACAAATTTTAAAACATTTCCACGTAATTTTAAATTTCTCTTCCAATATTTTAANTATTATG 2848
Db	308	ACAAATTTTAAAACATTTCCACATATTTTAAATTTATGGT-----ATACACTG 355
QY	2849	AAGCTCGAGTTTCAAAAGTGA---TTTTTTCCCACAAAGTGCCAACTTAAGCTAGAG 2905
Db	356	AAGCTCGAGTTTCAAAAGTGAATTTTTTTCCCACAAAGTTTCCACACACTTAAGCTAGAA 415
QY	2906	CTTTCAGTGTAAACTTTGCCCT-AAAAGTTAAGACATATTCTCAGAAATCATATATAGTCAC 2964
Db	416	CTTTCAGTGTAACTTTGCCCTAAAAGTTAAGACATATTTCTCATATCATACACAGTCAC 475
QY	2965	ATGATTTCTTGAGCTATCTGCTGTGTAAATAACAAAGATTCA 3007

Db 476 ATGATTTCTGATGCTATCTGGCCGTGTTTAATAATAAGATCTTTTA 518

RESULT 12  
 AI684386 548 bp mRNA linear EST 13-DEC-1999  
 LOCUS t696n02.x1 NCI\_CGAP\_CLL1 Homo sapiens cDNA clone IMAGE:2074035 3,  
 DEFINITION mRNA sequence.  
 ACCESSION AI684386  
 VERSION AI684386.1 GI:4895680  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 548)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 TITLE Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: Ash Alizadeh, John Byrd, M.D., Mike Grever,  
 M.D., Louis M. Staudt, M.D., Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D.  
 cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
 Insert length: 707 Std Error: 0.00  
 Seq primer: -400P from Gibco  
 High quality sequence stop: 471.  
 Location/Qualifiers  
 source  
 1..548  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2074035"  
 /clone\_lib="NCI\_CGAP\_CLL1"  
 /tissue\_type="B-cell, chronic lymphocytic leukemia"  
 /lab\_host="DH10B"  
 /note="Vector: p7T3D-Pac (Pharmacia) with a modified  
 polylinker; Site\_1: Not I; Site\_2: Eco RI; lstr strand cDNA  
 was primed with a Not I - oligo(dT) primer [5'  
 TGTACATCTGAAGTGGAGCGCGCATGCTGTTTTTTTTTTTTTTTTTTTTT  
 T 3']; double-stranded cDNA was ligated to Eco RI  
 adaptors (Pharmacia), digested with Not I and cloned into  
 the Not I and Eco RI sites of the modified p7T3 vector.  
 Library is normalized, and was constructed by Bento  
 Soares and M.Fatima Bonaldo."  
 BASE COUNT 214 a 91 c 64 g 179 t  
 ORIGIN  
 Query Match 11.5%; Score 358.2; DB 9; Length 548;  
 Best Local Similarity 88.3%; Pred. No. 2.3e-45;  
 Matches 455; Conservative 0; Mismatches 33; Indels 27; Gaps 5;  
 QY 2508 TTACAGCTTATAACACAACTTTTATTAGAAAAGTTATACATAACAGCATCAACTATTT 2567  
 Db 1 TTACAGCTTATAACACAACTTTTATTAGAAAAGTTATACATAACATAGCATCAACTATTT 60  
 QY 2568 TCAAG-----ACCCCAATAAGCAACAAAACAGCACTACAAAATGTGTAAACAAG 2617  
 Db 61 TCAAGAACAAATATTAACCCCGATTAAGCAACAAAACAGCACTACAAAATGTGTAAACAAG 120  
 QY 2618 AACTAATGACCTTCTTAAATCAAAACATTCAAATTAFTCTACAATGCTATTATCAACAACAG 2677  
 Db 121 AACTAATGACCTTCTTAAATCAACATTCAAATTAFTCTACAATGCTATTATCAACAACAG 180  
 QY 2678 GGAATACTCATGGTTTACAGGCATGTCAATATTGAAAATAAAGCTGCAATAGC-TTTTTTA 2736  
 Db 181 GGAATACTCTTGGTTTACAGGCATCAATATTGAATATAAAGCTGCAATAGCAATTTTA 240

Qy 2737 TACAATTATCGCTCTCAAGAAAATGAATCATTAAGACAGTAATTAGGAGTTTCACAAATTT 2796  
 Db 241 TACAATTACCACACTCTGAAGAACTGAATCATTAACACAGTAATTAAGAGTTTCACAAATTT 300  
 Qy 2797 AAAACATTTCCAGCTAATTTTAAATTTATGCTTCAATAATTTTAAATTTATTTGAAGTCTGA 2856  
 Db 301 AAAACATTTCCAGCTAATTTTAAATTTATGCTTCAATAATTTTAAATTTATTTGAAGTCTGA 348  
 Qy 2857 GTTTCAAAAGTGA---TTTTCCTCCCAAAAGGTGCCAACACTTTAAGCTAGAGCTTTTCAGT 2913  
 Db 349 GTTTCAAAAGTGAATTTTTCCTCCCAAAAGTTTCAACACTTTAAGCTAGAGCTTTTCAGT 408  
 Qy 2914 GTTAACCTTTGCCCT-AAAAGTTAAGACATATTTCTGAGATCATTAATAGTCACATGATTTTC 2972  
 Db 409 GTTAACCTTTGCCCTTAAAGTCTAAGACATATTTCTGATATCATCAACAGTCACATGATTTTC 468  
 Qy 2973 TGATGCTATCTGCTCTGTTAATACAAAGATTTCA 3007  
 Db 469 TGATGCTATCTGCTCTGTTAATATAAGTCTTTA 503

BM968219 577 bp mRNA linear EST 21-MAR-2002  
 LOCUS UI-CF-DU1-aa0-i-14-0-UI.s1 UI-CF-DU1 Homo sapiens cDNA clone  
 DEFINITION UI-CF-DU1-aa0-i-14-0-UI 3', mRNA sequence.  
 ACCESSION BM968219  
 VERSION BM968219.1 GI:195985806  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 577)  
 REFERENCE Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 AUTHORS Normalization and subtraction: two approaches to facilitate gene  
 TITLE discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 COMMENT Contact: McCray, PB  
 University of Iowa  
 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA  
 Tel: 319 356 4866  
 Fax: 319 356 7171  
 Email: paul-mccray@uiowa.edu  
 Tissue Procurement: Dr. M. J. Welsh, University of Iowa  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research  
 Genetics (www.resgen.com).  
 Seq primer: M13 FORWARD  
 POLYA=Yes.

FEATURES  
 source Location/Qualifiers  
 1..577  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="UI-CF-DU1-aa0-i-14-0-UI"  
 /clone\_lib="UI-CF-DU1"  
 /tissue\_type="Primary Lung Epithelial Cells"  
 /dev\_stage="Adult"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a  
 modified polylinker; Site\_1: EcoR I; Site\_2: Not I;  
 UI-CF-DU1 is a normalized cDNA library containing the  
 following tissue(s): Primary Lung Epithelial Cells The  
 library was constructed according to Bonaldo, Lennon and  
 Soares, Genome Research, 6:791-806, 1996. First strand  
 cDNA synthesis was primed with an oligo-dT primer  
 containing a Not I site. Double stranded cDNA was ligated  
 to an EcoR I adaptor, digested with Not I, and cloned  
 directionally into pT73-Pac vector. The oligonucleotide

used to prime the synthesis of first-strand cDNA contains  
 a library tag sequence that is located between the Not I  
 site and the (dri)18 tail. The sequence tag for this  
 library is GGCTGTAGGC.

TAG\_L1B=UI-CF-DU1  
 TAG\_TISSUE=Lung Epithelial Cells Tissue nos 359-368  
 TAG\_SEQ=GGCTGTAGGC"

BASE COUNT 213 a 71 g 196 t  
 ORIGIN

Query Match 11.5%; Score 358.2; DB 14; Length 577;  
 Best Local Similarity 87.6%; Pred. NO. 2.3e-45;  
 Matches 458; Conservative 0; Mismatches 38; Indels 27; Gaps 5;

Qy 2500 TATTTTTCAGCTTATAACACACACTTTTATTAGAAAAGTTATACATACACAGCATC 2559  
 Db 9 TTTTTCAGCTTATAACACACACTTTTATTAGAAAAGTTATACATACACAGCATC 68  
 Qy 2560 AACTATTTTCAAG-----AACCCCAATAGCAACAAACACAGACTAACAAAATGT 2609  
 Db 69 AACTATTTTCAAGACAATATTAAACCCGATAAGCAACAAACACAGACTAACAAAATGT 128  
 Qy 2610 GTAACAAGAACTAATGACCTTTCTAAAATCAACATTCATATATCTACATGTCTATT 2669  
 Db 129 GTAACAAGAACTAATGACCTTTCTAAAATCAACATTCATATATCTACATGTCTATT 188  
 Qy 2670 ACAACAGGGAACCTCCATGTTTACAGGCATGTCATATTCAAAATAAGCTGCAATAG 2729  
 Db 189 ACAACGGGGAACCTCCATGTTTACAGGCATCATATTAATATAAGCTGCAATAG 248  
 Qy 2730 C-TTTTATACAAATTAATCGCTCTCAAGAAATGAATCATTAAGACAGTAATTTAGGAGTTC 2788  
 Db 249 CAATTTATACAAATTAATCGCTCTCAAGAAATGAATCATTAAGACAGTAATTTAGGAGTTC 308  
 Qy 2789 ACAATTTAAACATTTTCAGTAATTTTAAATATTGCTTCAATAATTTTAAATATTG 2848  
 Db 309 ACAATTTAAACATTTTCAGTAATTTTAAATATTGCTTCAATAATTTTAAATATTG 356  
 Qy 2849 AGCTCGAGTTTCAAAAGTGA---TTTTTCCCACAAAGGTGCCAACACTTAAGCTAGAG 2905  
 Db 357 AGCTCGAGTTTCAAGAGTGAATTTTTCACACAAAGTTTTCACACTTAAGCTAGAA 416  
 Qy 2906 CTTTCAGTGTTAACCTTTGCCCT-AAAAGTTAAGACATATTTCTGAGAAATCATATAGTCAC 2964  
 Db 417 CTTTCAGTGTTAACCTTTGCCCTAAAAAGTTAAGACATATTTCTGATAATCATACAGTCAC 476  
 Qy 2965 ATGATTTCTGATGCTATCTGCTCTGTGTTAATAACAAGATTTCA 3007  
 Db 477 ATGATTTCTGATGCTATCTGCTCTGTGTTAATAACAAGATTTCA 519

## RESULT 14

BM715473

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

BM715473 725 bp mRNA linear EST 28-FEB-2002  
 UI-E-EJ0-ahi-k-10-0-UI.r2 UI-E-EJ0 Homo sapiens cDNA clone  
 UI-E-EJ0-ahi-k-10-0-UI 5', mRNA sequence.

BM715473

BM715473.1 GI:19028731

EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 725)

Bonaldo,M.F., Lennon,G. and Soares,M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

97044477

Contact: Soares, MB

Program for Rat Gene Discovery and Mapping

University of Iowa

451 Eckstein Medical Research Building Iowa City, IA 52242, USA

Tel: 319 335 8250  
Fax: 319 335 9565  
Email: msoares@blue.weeg.uiowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics (www.resgen.com).  
The following repetitive elements were found in this cDNA  
sequence: 12-40, >POLY\_A\$imple\_repeat (matched complement)  
Seq primer: M13 Reverse.

FEATURES  
source  
Location/Qualifiers  
1..725  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="UI-E-EJ0-ah1-k-10-0-UI"  
/clone\_lib="UI-E-EJ0"  
/tissue\_type="fetal eyes, lens, eye anterior segment, optic nerve, retina, Retina Foveal and Macular, RPE and Choroid"  
/dev\_stage="fetal and adult"  
/lab\_host="PH10B (Life Technologies) (T1 phage resistant)"  
/note="Organ: eye; Vector: pT7m3-Pac (Pharmacia) with a modified polylinker; Site\_1: EcoR I; Site\_2: Not I; UI-E-EJ0 is a subcloned cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7m3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequencing tags for this library are: fetal eyes, AGAATCAAGA; lens, CGATTACGGA; eye anterior segment, AATGCCGAT; optic nerve, CCATTAACTG; retina, CCGCG; Retina Foveal and Macular, GTCC; RPE and Choroid, ACTCA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NET)."  
BASE COUNT 262 a 122 c 97 g 242 t 2 others

Query Match 11.4%; Score 355.6; DB 14; Length 725;  
Best Local Similarity 87.4%; Pred. No. 5e-45; Mismatches 0; Indels 27; Gaps 5;  
Matches 456; Conservative 0;  
QY 2500 TATTTTTTTTACAGCTTATACACAACTTTTATTAGAAAAGTTATACATACACAGCATC 2559  
DB 31 TTTTTTTTTTACAGCTTATACACAACTTTTATTAGAAAAGTTATACATACATAGCATC 90  
QY 2560 AACTATTTTCAAG-----ACCCAAATAGCAACAAAACCCAGACTAACAAAATGT 2609  
DB 91 AACTATTTTCAAGAAACAATATTAAACCCCGATAGCAACAAAACCCAGACTAACAAAATGT 150  
QY 2610 GTAAACAGAAACTAATGACCTTCTTAAATCAACATTCATTAATCTACATGCTATT 2669  
DB 151 GTAAACAGAAACTAATGACCTTCTTAAATCAACATTCATTAATCTACATGCTTTT 210  
QY 2670 ACAACAGGGAACATCCATGCTTTACAGGCATGTCTATATTGAAATAAAGCTGCAATAG 2729  
DB 211 ACAACAGGGAACATCCCTTGGTTTACAGGCACATCATATTGAATATAAAGCTGCAATAG 270  
QY 2730 C-TTTTATACAAATATCGCTCTCAAGAAATGAATCATTTAAGACAGTAATAGAGTTTC 2788  
DB 271 CAATTTTATACAAATATCCACTCTGAAGAACTGAATCATTTAAGACAGTAATAGAGTTTC 330  
QY 2789 ACAAAATTTAAACATTTACAGTAAATTTTAAATTTATTGCTTCAATAATTTTAAATTTTG 2848  
DB 331 ACAAAATTTAAACATTTACATAATTTTAAATTTATTGGGT-----ATACACTG 378  
QY 2849 AGCTCTGAGTTTCAAAAGTGA---TTTTTTTCCCAAAAGGTGCCAACACTTTAAGCTAGAG 2905  
|||||

DB 379 AAGTCTGAGTTTCAAAAGTGAATTTTTTTTCCACAAAAGTTTCAACACTTAAAGCTAGAA 438  
QY 2906 CTTTCAGTGTAACTTTGCCCTAAAGAGTTAAAGACATATTTCTCAGAAATCATATAGTCACA 2965  
|||||  
DB 439 CTTTCAGTGTAACTTTGCCCTAAAGAGTTA-AGACATTCTGATAATCATACAGTCACA 497  
QY 2966 TGATTTCATGCTATCTGCTCTGTGTTAATAACAAGATTTC 3007  
DB 498 TGATTTCTGATGCTATCTGCTGTTGTTAATAATAAGTCTTTA 539  
RESULT 15  
AA557129 496 bp mRNA linear EST 09-SEP-1997  
LOCUS nl74a09.sl NCI\_CGAP\_Br2 Homo sapiens cDNA clone IMAGE:1056376 3',  
DEFINITION mRNA sequence.  
ACCESSION AA557129  
VERSION AA557129.1 GI:2327606  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 496)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-f@mail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/bbrp/image/image.html  
Insert length: 656 Std Error: 0.00  
Seq primer: -40ml3 fwd. ET from Amersham  
High quality sequence stop: 474.  
location/Qualifiers  
1..496  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1056376"  
/clone\_lib="NCI\_CGAP\_Br2"  
/sex="female, pooled"  
/tissue\_type="breast"  
/lab\_host="DH10B"

/note="Vector: pT7m3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from pooled bulk breast tumor tissue, and was then primed with a Not I-oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7m3 vector. This library is the normalized version of NCI\_CGAP\_Br1.1. Library was constructed by Bento Soares and M. Fatima Bonaldo."  
BASE COUNT 198 a 86 c 57 g 155 t  
ORIGIN  
Query Match 11.4%; Score 355; DB 9; Length 496;  
Best Local Similarity 88.8%; Pred. No. 7.5e-45;  
Matches 450; Conservative 0; Mismatches 30; Indels 27; Gaps 5;  
QY 2510 ACAGCTTATACACAACTTTTATTAGAAAAGTTATACATACACAGCATCACTATTTC 2569  
DB 1 ACAGCTTATACACAACTTTTATTAGAAAAGTTATACATACATAGCATCACTATTTC 60  
QY 2570 AAG-----AACCCAAATAGCAACAAAACCCAGACTAACAAAATGTGTAACAAGAA 2619  
DB 61 AAGAACAATATTAAACCCGTAAGCAACAAAACCCAGACTAACAAAATGTGTAACAAGAA 120  
|||||



Qy	2620	ACTAATGACCTTTCTAAAATCAACACATTCAAATTCTCAATGCTCTATTTCACAAACAGGG	2679
Db	121	ACTAATGACCTTTCTAAAATCAACACATTCAAATTCTCAATGCTCTTTTACAAACCGGG	180
Qy	2680	AAACATCCATGGTTTACAGGCATGTCATATTCAAATAAAGCTGCAATAGC-TTTTTTATA	2738
Db	181	AAACATCCTTGGTTTACAGGCACATCATATTGAATATAAAGCTGCAATAGCAATTTTATA	2404
Qy	2739	CAATTATCGCTCTCAAGAAAAATGAATCATTAAGACAGTAAATTAGGAGTTCCACAAATTTAA	2798
Db	241	CAATTACCCTCTGGAAGAAACTGAATCATTAAGACAGTAAATTACGAGTTCCACAAATTTAA	300
Qy	2799	AACATTTACGTAATTTTAAATATTGCTTCCAATAATTTTAAATATTGAAGTCTGAGT	2858
Db	301	AACATTTACATAATTTTAAATATTGCTTCCAATAATTTTAAATATTGAAGTCTGAGT	348
Qy	2859	TTCAAAAGTCA---TTTTTTCCCAAAAGGTCCCAACACTTAAAGCTAGAGCTTTTCAGTGT	2915
Db	349	TTCAAAAGTGAATTTTTTTTCCCAAAAGTTTCAACACTTAAAGCTAGAGCTTTTCAGTGT	408
Qy	2916	TAACTTTTGCCT-AAAAGTTAAGACATATTTCTGAGAATCATATAAGTACACATGATTTCTG	2974
Db	409	TAACTTTTGCCTAAAAGTTAAGACATATTTCTGATATAATCAACAGTCACATGATTTCTG	468
Qy	2975	ATGCTATCTGCTCTGTTTAATACAAAG	3001
Db	469	ATGCTATCTGCTCTGTTTAATATAAAG	495

Search completed: February 17, 2003, 11:47:10  
Job time : 3022.64 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: February 17, 2003, 08:46:11 ; Search time 5370.63 Seconds  
(without alignments)  
16874.395 Million cell updates/sec  
Title: US-09-766-511B-51  
Perfect score: 3114  
Sequence: 1 cttaatgttggaagtctctt.....tttaaaaaaaaaaaaaaaaaaaaa 3114  
Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :				GenEmbl:*			
	1:	gb_ba:*					
	2:	gb_htg:*					
	3:	gb_in:*					
	4:	gb_om:*					
	5:	gb_ov:*					
	6:	gb_pat:*					
	7:	gb_ph:*					
	8:	gb_pl:*					
	9:	gb_pr:*					
	10:	gb_ro:*					
	11:	gb_sts:*					
	12:	gb_sy:*					
	13:	gb_un:*					
	14:	gb_vi:*					
	15:	em_ba:*					
	16:	em_fun:*					
	17:	em_hum:*					
	18:	em_in:*					
	19:	em_mu:*					
	20:	em_om:*					
	21:	em_or:*					
	22:	em_ov:*					
	23:	em_pat:*					
	24:	em_ph:*					
	25:	em_pi:*					
	26:	em_ro:*					
	27:	em_sts:*					
	28:	em_un:*					
	29:	em_vi:*					
	30:	em_htg_hum:*					
	31:	em_htg_inv:*					
	32:	em_htg_other:*					
	33:	em_htg_mus:*					
	34:	em_htg_pln:*					
	35:	em_htg_rod:*					
	36:	em_htg_man:*					
	37:	em_htg_vrt:*					
	38:	em_sy:*					
	39:	em_htgo_hum:*					
	40:	em_htgo_mus:*					
	41:	em_htgo_other:*					

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
C 1	2446.6	78.6	109320	9	AC092746	AC092746 Homo sapi
C 2	379.2	12.2	1227	6	AX155225	AX155225 Sequence
C 3	379.2	12.2	1227	10	AF240357	AF240357 Mus muscu
C 4	379.2	12.2	1254	10	BC023008	BC023008 Mus muscu
C 5	364.6	11.7	3708	6	AX014883	AX014883 Sequence
C 6	357.2	11.5	156441	9	AP004074	AP004074 Homo sapi
C 7	356.2	11.4	2209	9	AK097915	AK097915 Homo sapi
C 8	356.2	11.4	3647	9	HUMORF005	D38551 Human mRNA
C 9	346.6	11.1	193100	2	AC016729	AC016729 Homo sapi
C 10	346.6	11.1	193364	9	AC087350	AC087350 Homo sapi
C 11	317.6	10.2	37652	9	HS085A3	Z78021 Human DNA s
C 12	317.6	10.2	56583	9	HS347M6	295327 Human DNA s
C 13	283.8	9.1	1312	6	AX155223	AX155223 Sequence
C 14	283.8	9.1	1313	9	AF293615	AF293615 Homo sapi
C 15	280.6	9.0	851	9	AF325459	AF325459 Homo sapi
C 16	276.8	8.9	1125	10	AF240358	AF240358 Mus muscu
C 17	272.8	8.8	827	6	AX357481	AX357481 Sequence
C 18	269	8.6	800	6	AX357483	AX357483 Structural
C 19	257.8	8.3	3298	6	E11759	E11759 Structural
C 20	257.8	8.3	3298	10	MUSB14GA	D37790 Mouse mRNA
C 21	255.4	8.2	684	6	AX098173	AX098173 Sequence
C 22	244	7.8	1125	10	AF240359	AF240359 Mus muscu
C 23	235	7.5	428	11	G23519	G23519 human STS W
C 24	233.2	7.5	739	9	AF325460	AF325460 Homo sapi
C 25	224.4	7.2	3929	6	AX305440	AX305440 Sequence
C 26	224.4	7.2	3929	10	MUSF9	D49429 Mouse NCBP-
C 27	201.2	6.5	444	6	AX456975	AX456975 Sequence
C 28	200.4	6.4	402	6	AX456962	AX456962 Sequence
C 29	197	6.3	1133	9	AF109146	AF109146 Homo sapi
C 30	197	6.3	1195	6	AX400064	AX400064 Sequence
C 31	197	6.3	1271	9	HS4133532	AJ133532 Homo sapi
C 32	195.4	6.3	1305	9	AF067800	AF067800 Homo sapi
C 33	194.2	6.2	615	9	AF200738	AF200738 Homo sapi
C 34	186	6.0	1308	9	AF328684	AF328684 Homo sapi
C 35	182.4	5.9	55577	2	AC107915	AC107915 Homo sapi
C 36	182.4	5.9	117296	9	AC092865	AC092865 Homo sapi
C 37	179.4	5.8	582	6	AX456965	AX456965 Sequence
C 38	177.8	5.7	753	6	AX456967	AX456967 Sequence
C 39	177.6	5.7	110000	2	AC092450_2	Continuation (3 of
C 40	169.6	5.4	1013	6	AX456966	AX456966 Sequence
C 41	167.6	5.4	176924	2	AC111632	AC111632 Rattus no
C 42	166.2	5.3	125818	2	AC098495	AC098495 Rattus no
C 43	154.8	5.0	558	6	AX456968	AX456968 Sequence
C 44	154.8	5.0	1192	10	BC034893	BC034893 Mus muscu
C 45	148	4.8	55577	2	AC107915	AC107915 Homo sapi

ALIGNMENTS

RESULT 1  
AC092746/c  
LOCUS AC092746 109320 bp DNA linear PRI 12-JUN-2002  
DEFINITION Homo sapiens 12 BAC RP11-561P12 (Roswell Park Cancer Institute  
Human BAC Library) complete sequence.  
ACCESSION AC092746  
VERSION AC092746.9 GI:21392435  
KEYWORDS HTG.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 109320)  
AUTHORS Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,  
Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T.,

Barbaria, J., Benton, J., Himage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Caron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chiu, D., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Emerling, S., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J.J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Han, J., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Honsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Ioshikhes, I., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lee, E., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulsegod, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Marondel, I., Martin, R., Martindale, A., Martinez, E., Massey, E., Mathoney, E., McLeod, M.P., Meador, M., Mei, G., Merscher, S., Metzker, M., Miller, A., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Montgomery, K.T., Morgan, M., Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shim, C., Shoshitari, N., Sisson, I., Sodergren, E., Sonaite, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usman, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Kucherlapati, R., Weinstock, G. and Gibbs, R.

Direct Submission  
Unpublished  
2 (bases 1 to 109320)  
Worley, K.C.  
Direct Submission  
Submitted (25-JUL-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 109320)  
Worley, K.C.  
Direct Submission  
Submitted (07-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
4 (bases 1 to 109320)  
Worley, K.C.  
Direct Submission  
Submitted (12-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Jun 12, 2002 this sequence version replaced gi:21217374.  
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email [gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

#### ANNOTATION OF FEATURES:

STs are identified using ePCR (Genome Res. 7:541-550) searches

of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

FEATURES	source
	Location/Qualifiers
	1..109320
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="12"
	/clone="RP11-561P12"
misc_feature	1..2004
	/note="overlaps bases 69129..71132 of clone AC092184"
	/function="clone overlap"
repeat_region	787..1014
	/rpt_family="MIR"
repeat_region	complement(1235..1452)
	/rpt_family="MIR"
repeat_region	1686..1789
	/rpt_family="L2"
repeat_region	1927..2209
	/rpt_family="AluSq"
repeat_region	2210..2238
	/rpt_family="(TAAAA)n"
repeat_region	2430..2741
	/rpt_family="AluJo"
repeat_region	2938..3085
	/rpt_family="L2"
repeat_region	complement(3086..3380)
	/rpt_family="AluSg"
repeat_region	3381..3665
	/rpt_family="L2"
repeat_region	3698..3940
	/rpt_family="AluJb"
repeat_region	4033..4297
	/rpt_family="AluSx"
repeat_region	4366..4397
	/rpt_family="(TTTA)n"
repeat_region	complement(4398..4689)
	/rpt_family="AluJb"
repeat_region	complement(4706..4993)
	/rpt_family="AluSx"
repeat_region	5253..5592
	/rpt_family="AluSx"
repeat_region	complement(5611..5872)
	/rpt_family="L2"
repeat_region	6027..6315
	/rpt_family="MLTIF"
repeat_region	complement(6371..6498)
	/rpt_family="MIR"
repeat_region	6703..7006
	/rpt_family="AluSx"
repeat_region	complement(7277..7397)
	/rpt_family="FLAN_C"



Qy	2135	TGAGGTATATACAATGTTATTAATTTATCTACTGTGTCATCGGCAAAATTAATTTTTTATCATTT	2184
Db	9892	CGAGTATATACAATGTTATTAATTTATCTACTGTGTCATCGGCAAAATTAATTTTTTATCATTT	98933
Qy	2195	GTGAGTGCACATTTCTTTAAATTTAGTAATGCTTTGGCTTTAAATTTTCTCCTCATATTA	2254
Db	9893	GTGAGTGCACATTTCTTTAAATTTAGTAATGCTTTGGCTTTAAATTTTCTCCTCATATTA	99773
Qy	2255	AAATAGATACAGTAACATTTTCATTTATGTAGTGTGCTGTAATAATTTTTTTTCCATCTTCTAT	2314
Db	9972	AAATAGATACAGTAACATTTTCATTTATGTAGTGTGCTGTAATAATTTTTTTTCCATCTTCTAT	99713
Qy	2315	TTTTGACCATTTTTTATTTCCACATGTGCTCTTAATAAGTAGCATATAGTTAAATTTTAAAA	2374
Db	9971	TTTTGACCATTTTTTATTTCCACATGTGCTCTTAATAAGTAGCATATAGTTAAATTTTAAAA	99653
Qy	2375	AATCCAATATGGCAATCACCTTTTAGGTTAAAAATTTAATCCATTTACATTTGTGACAAT	2434
Db	9965	AATCCAATATGGCAATCACCTTTTAGGTTAAAAATTTAATCCATTTACATTTGTGACAAT	99593
Qy	2435	TCGACATATATATGGTCTTAAATCTATCATCTTACTTAGTGGTGGTTCCATTTCCCTCTGCTC	2494
Db	9959	TCGACATATATATGGTCTTAAATCTATCATCTTACTTAGTGGTGGTTCCATTTCCCTCTGCTC	99533
Qy	2495	CAAAATATTTTTTTTACAGCTTTATAACACACACTTTTTATTAGAAAAGTTATACATAACACA	2554
Db	9953	CAAAATATTTTTTTTACAGCTTTATAACACACACTTTTTATTAGAAAAGTTATACATAACACA	99473
Qy	2555	GCATCAACTATTTTCAAGAACCCCAATAGCAACAAAAACACAGACTAACAAAAATGTGAAC	2614
Db	9947	GCATCAACTATTTTCAAGAACCCCAATAGCAACAAAAACACAGACTAACAAAAATGTGAAC	99413
Qy	2615	AAGAACTAATGACCTTTCTTAAATTCAAACATTCCAATTTATCTACAATGTCTATTTACAAA	2674
Db	9941	AAGAACTAATGACCTTTCTTAAATTCAAACATTCCAATTTATCTACAATGTCTATTTACAAA	99353
Qy	2675	CAGGAAAACTCCATGTTTTACAGCATGTGCATATTTGAAAAATAAGCTGCAATAGCTTTT	2734
Db	9935	CAGGAAAACTCCATGTTTTACAGCATGTGCATATTTGAAAAATAAGCTGCAATAGCTTTT	99293
Qy	2735	TATCAATATTCGCTCTCAAGAAAAATGAATCAATTAAGACAGATAATTAGGAGTTCCACAAAT	2794
Db	9929	TATCAATATTCGCTCTCAAGAAAAATGAATCAATTAAGACAGATAATTAGGAGTTCCACAAAT	99233
Qy	2795	TTAAACATTTTCCGTAATTTTAAATTTATCTCTTCAATAATTTTAAATTTATTCGAAGTCT	2854
Db	9923	TTAAACATTTTCCGTAATTTTAAATTTATCTCTTCAATAATTTTAAATTTATTCGAAGTCT	99173
Qy	2855	GAGTTTCAAAAGTGA - TTTTTTCCCAAGAGGTGCCAACACTTAAGCTAGAGCTTTTCAGT	2913
Db	9917	GAGTTTCAAAAGTGAATTTTTTCCCAAGAGGTGCCAACACTTAAGCTAGAGCTTTTCAGT	99113
Qy	2914	GTTAACTTTGGCCTTAAAGTTTAAGACATATTTCTGAGAAATCATATAGTCACATGATTTCT	2973
Db	9911	GTTAACTTTGGCCTTAAAGTTTAAGACATATTTCTGAGAAATCATATAGTCACATGATTTCT	99053
Qy	2974	GATGCTATCTGCTGTTTAATAACAAAGATTTTCACACATGAATACCTATGTAAACAAATCT	3033
Db	9905	GATGCTATCTGCTGTTTAATAACAAAGATTTTCACACATGAATACCTATGTAAACAAATCT	98993
Qy	3034	CCATGTTCTACACATATATCCCGAGAACTTAAAGTATATAATAATAAACATAGCAAGC	3093
Db	9899	CCATGTTCTACACATATATCCCGAGAACTTAAAGTATATAATAATAAACATAGCAAGC	98933
Qy	3094	CTTTA 3098	
Db	9893	CTTTA 98928	
RESULT 2			
AX155225			
LOCUS	AX155225	1227 bp	DNA
DEFINITION	Sequence 3 from Patent WO0136487.	linear	PAT 22-JUN-2001

## RESULT 2

AX155225

LOCUS

### DEFINITION



Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowls, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: iRAK Plate: 53 Row: p Column: 4  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 9910157.

FEATURES

source  
1..1254  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/map="FVB/N-3"  
/clone="MGC:35901 IMAGE:5376247"  
/tissue\_type="Mammary tumor. MMTV-LTR/INT3 model. 5 month old mouse. Taken by biopsy."  
/clone\_lib="NCI-CGAP\_Mam2"  
/lab\_host="DH10B"  
/note="vector: pCMV-SPORT6"  
146..775  
/codon\_start=1  
/product="C-type (calcium dependent, carbohydrate recognition domain) lectin, superfamily member 10"  
/protein\_id="AAH23008.1"  
/db\_xref="GI:18606438"  
/db\_xref="LOCUSID:56620"  
/translation="MVGEROSGKVGCTLRLSAAVISMILLSTCFIASCVVYQFI MDPSRLYELHYTHSLTFCSEGTVMYSEKMGCCPNHKSFGSCYLLSTKFNWST SEQNCVOMGHLVINTAEQONFTIOLNLSLYFLGLSDPQNGKQWIDDTFFSQN VRFVHEPNLPERCQSVIYVNPWSKMGVDFCDSKHNSICEMKKIYL"

CDS

362 a 271 c 276 g 345 t  
BASE COUNT  
ORIGIN  
Query Match 12.2%; Score 379.2; DB 10; Length 1254;  
Best Local Similarity 73.8%; Pred. No. 1.2e-53;  
Matches 555; Conservative 0; Mismatches 178; Indels 19; Gaps 5;

QY 40 GCAGTTTGTCCCTGAGCTCTAGCTCTTTAAATGAAGCTGAGTCTCTGGGCAACATCTTT 99  
DB 39 GGAAGTTGATTCTGAACCTGGCTCTTTGACAGAGCCAGTCCCTGAGTCGTATTTT- 97  
QY 100 AGGGAGAGAGGTACAAAAGTTTCCCTGACCTTCTCAACACAGAGGAGCCTGCATGATG 159  
DB 98 --GGAGACAGATGCAAGAAACCCCT-GACCTTCTGAACATA---CACCTCAACAATGGTG 151  
QY 160 CAAGAGCAGCAACCTCAAGTACAGAAAGAGGCTGGTTGTCCTCGAGACTCTGGTCT 219  
DB 152 CAGGAAGACAAATCCCA-----AGGGAAGGAGTCTGCTGGACCTGAGACTCTGGTCA 205  
QY 220 GTGGCTGGGATTTCATGCGACTCCTCAGTCTGCTTCATTGTGAGCTGTGTAGTAAC 279  
DB 206 GCTGCTGTGATTTCATGTTACTCTGTAGTACCTGTTTCATTTGGAGCTGTGTGACT 265  
QY 280 TACCAATTTACATATGGTGAACACTGGCAAAAGGCTGTCTGAACCTACACTCATATCA 339  
DB 266 TACCAATTTATATGGACAGCCAGTGAAGACTATATGAACCTTCACACATACCATTC 325  
QY 340 AGTCTCACTGCTTCAGTGAAGGCAAGGTGCCAG-----CCTGGGATGTGGCCA 393  
DB 326 AGTCACACTGCTTCAGTGAAGGAGTATGGTGTGCAGAAAATGTGGGGAIGTGGCCA 385  
QY 394 GCTTCTTGAAGTCAATTTGGTTCCAGTTGCTACTTTCATTTCCAGTGAAGAGGTTTG 453  
DB 386 AATCACTGGAAGTCAATTTGGTCCAGCTGCTACCTCATTTCTACCAAGGAGAACTTC 445  
QY 454 TCTAAGATGACAGAACTGTGTGAGATGGAGACATATTTGGTTGTGTTTCAACACAGAA 513  
DB 446 AGCACAGTGGACAGAACTGTGTTTCAGATGGGGCTCATCTGGTGGTGATCAATACTGAA 505  
QY 514 GCAGAGCAGAAATTTCAATTCACAGAGCTGAATGATCATTTTCTTATTTTCTGGGCTT 573  
DB 506 GCGGAGCAGAAATTTCAATCCAGCAGCTGAATGAGTCACCTTTCTTACTTCTGGGCTT 565

QY 574 TCAGACCCACACAGGTAAATAATTTGGCAATGGATTGATGAACACACCTTATGAGAAAAT 633  
DB 566 TCGGATCCACACAGGTAAATGCAAAATGGCAATGGATGATGATGATGATGATGATGATG 625  
QY 634 GTCAGATTTTGGCACCTAGGTGAGCCCAATCTCTGCAGAGCAATGTCTTCAATATGTC 693  
DB 626 GTCAGTTCTGGCAGCCCAATGAAATCTTCCAGAGAGGGGTGTGTTTCAATATGTT 685  
QY 694 TTTCTGAAAACCTACAGGATGGGGCTGGAATGATGTTTATCTGTGAAACTAGAAGGAATCA 753  
DB 686 TACTGGAATCTCGAAATGGGCTGGAATGATGTTTCTGTGATAGTAACACAAATCA 745  
QY 754 ATATGTGAGATGAATAAGATTTACCTATGAGT 785  
DB 746 ATATGTGAAATGAAGAAGATTTACCTATGAGT 777  
RESULT 5  
AX014883 LOCUS 3708 bp DNA linear PAT 07-SEP-2000  
DEFINITION Sequence 81 from Patent WO9953040.  
ACCESSION AX014883  
VERSION AX014883.1 GI:10041150  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 3708)  
AUTHORS Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and Pilarsky,C.  
TITLE Human nucleic acid sequences from ovarian tumour tissue  
JOURNAL Patent: WO 9953040-A 81 21-OCT-1999;  
SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN (DE); PILARSKY CHRISTIAN (DE)  
FEATURES  
Location/Qualifiers  
source 1..3708  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 1036 a 818 c 671 g 1183 t  
ORIGIN  
Query Match 11.7%; Score 364.6; DB 6; Length 3708;  
Best Local Similarity 88.3%; Pred. No. 2.7e-51;  
Matches 462; Conservative 0; Mismatches 34; Indels 27; Gaps 5;



```

|||||
Db 388 AAGTCTGAGTTCAAAAGTCATTTTTCCTCCACAAAGTTTCACACTTAGCTAGAA 447
|||||
QY 2906 CTTTCAGTGTAACCTTGCCT-AAAAGTTAAGACATATTCGAGAAATCAATAGTCAC 2964
|||||
Db 448 CTTTCAGTGTAACCTTGCCTAAAAGTTAAGACATATTCGATATCAATACAGTCAC 507
|||||
QY 2965 ATGATTTCTGATCTATCTCTGTTGTTAATAACAAGATTTC 3007
|||||
Db 508 ATGATTTCTGATCTATCTCTGTTGTTAATAATAAAGTCITTA 550
|||||

RESULT 6
AP004074 156441 bp DNA linear PRI 21-AUG-2001
LOCUS Homo sapiens genomic DNA, chromosome 8q23, clone:KB1266F9, complete
DEFINITION
ACCESSION AP004074
VERSION AP004074.1 GI:15216353
KEYWORDS
SOURCE
Homo sapiens cell_line:FLEB 14-14 DNA, clone_lib:Keio BAC library
clone:KB1266F9.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Shimizu,N. and Asakawa,S.
AUTHORS
TITLE Homo sapiens DNA chromosome 8 SEQUENCE
JOURNAL Published Only in Database (2001)
REFERENCE
2 (bases 1 to 156441)
AUTHORS Shimizu,N. and Asakawa,S.
TITLE Direct Submission
JOURNAL Submitted (20-AUG-2001) Nobuyoshi Shimizu, Keio University, School
of Medicine, Molecular Biology; 35 Shinanomachi, Shinjuku-ku, Tokyo
160-8582, Japan (E-mail:nshimizue@med.keio.ac.jp,
Tel:81-3-3351-2370, Fax:81-3-3351-2370)
FEATURES
Source
1. 156441
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8q23"
/clone="KB1266F9"
/cell_line="FLEB 14-14"
/clone_lib="Keio BAC library"
560..581
/revidence-not_experimental
repeat_region /rpt_family="AT_rich"
2623..3033
/revidence-not_experimental
repeat_region /rpt_family="AT_rich"
3626..3688
/revidence-not_experimental
repeat_region /rpt_family="MSTB"
5506..5537
/revidence-not_experimental
repeat_region /rpt_family="(A)n"
7579..7632
/revidence-not_experimental
repeat_region /rpt_family="Ct-rich"
complement(7956..8247)
/revidence-not_experimental
repeat_region /rpt_family="AluSx"
8697..8731
/revidence-not_experimental
repeat_region /rpt_family="(CAAA)n"
9452..9560
/revidence-not_experimental
repeat_region /rpt_family="I2"
10358..10385
/revidence-not_experimental
repeat_region /rpt_family="AT_rich"
10421..10441
/revidence-not_experimental
/revidence-not_experimental
/rpt_family="AT_rich"
complement(10443..10592)
/revidence-not_experimental
/rpt_family="FLAM_C"
11163..11185
/revidence-not_experimental
/rpt_family="(A)n"
11849..11876
/revidence-not_experimental
/rpt_family="AT_rich"
12699..12861
/revidence-not_experimental
/rpt_family="FAM"
13403..13423
/revidence-not_experimental
/rpt_family="AT_rich"
13609..13719
/revidence-not_experimental
/rpt_family="MIR"
13731..13758
/revidence-not_experimental
/rpt_family="AT_rich"
13766..14028
/revidence-not_experimental
/rpt_family="L1MC5"
14034..14260
/revidence-not_experimental
/rpt_family="Tigger1"
complement(14263..14407)
/revidence-not_experimental
/rpt_family="FLAM_A"
14409..14877
/revidence-not_experimental
/rpt_family="Tigger1"
14872..15595
/revidence-not_experimental
/rpt_family="Tigger1"
15637..15922
/revidence-not_experimental
/rpt_family="L1MC5"
17431..17469
/revidence-not_experimental
/rpt_family="AT_rich"
19504..19526
/revidence-not_experimental
/rpt_family="AT_rich"
complement(20688..21125)
/revidence-not_experimental
/rpt_family="L1MC1"
complement(22612..22919)
/revidence-not_experimental
/rpt_family="L1MD1"
complement(22922..23099)
/revidence-not_experimental
/rpt_family="AluJo"
complement(23100..23442)
/revidence-not_experimental
/rpt_family="L1MD1"
23442..24177
/revidence-not_experimental
/rpt_family="L1MC/D"
24178..24474
/revidence-not_experimental
/rpt_family="AluY"
24475..24685
/revidence-not_experimental
/rpt_family="L1MC/D"
25120..25255
/revidence-not_experimental
/rpt_family="FLAM_C"
complement(26608..26781)
/revidence-not_experimental

```

```

repeat_region /rpt_family="MIR"
complement(26999..27171)
/evidence-not_experimental
/rpt_family="MER2"
repeat_region 27399..27410
/evidence-not_experimental
/rpt_family="AT_rich"
repeat_region 27411..27435
/evidence-not_experimental
/rpt_family="(CA)n"
repeat_region 29520..29564
/evidence-not_experimental
/rpt_family="G-rich"
repeat_region 30299..32109
/evidence-not_experimental
/rpt_family="LMA4"
repeat_region 33154..33345
/evidence-not_experimental
/rpt_family="MTR"
repeat_region 33496..33574
/evidence-not_experimental
/rpt_family="CT-rich"
repeat_region 34340..34598
/evidence-not_experimental
/rpt_family="AluSc"
repeat_region 34711..34754
/evidence-not_experimental
/rpt_family="AT_rich"
repeat_region 35069..35113
/evidence-not_experimental
/rpt_family="GA-rich"
repeat_region 35622..35733
/evidence-not_experimental
/rpt_family="FLAM_A"
repeat_region 35943..35972
/evidence-not_experimental
/rpt_family="AT_rich"
repeat_region 36516..36600
/evidence-not_experimental
/rpt_family="MER5A"
repeat_region 36507..36687
/evidence-not_experimental
/rpt_family="T-rich"
repeat_region 36731..36961
/evidence-not_experimental
/rpt_family="MIR"
repeat_region 37228
complement(36969..37228)
/evidence-not_experimental
/rpt_family="AluJb"
repeat_region 37373..37775
/evidence-not_experimental
/rpt_family="MLTJ"
repeat_region 38709..38978
/evidence-not_experimental
/rpt_family="L2"
repeat_region 39053..39073
/evidence-not_experimental
/rpt_family="AT_rich"
repeat_region 39582..39719
/evidence-not_experimental
/rpt_family="L2"
repeat_region 40014
complement(39720..40014)
/evidence-not_experimental
/rpt_family="LMB5"
repeat_region 40015..40420
/evidence-not_experimental
/rpt_family="MSTF"
repeat_region 40462
complement(40421..40462)

Query Match 11.5%; Score 357.2; DB 9; Length 156441;
Best Local Similarity 88.3%; Pred. No. 2.5e-50;
Matches 454; Conservative 0; Mismatches 33; Indels 27; Gaps 5;

QY 2509 TACAGCTTATACACAACTTTTATTAGAAAAGTTATATACACAGCATCAACTATTTT 2568
Db 798 TACAGCTTATACACAACTTTTATTAGAAAAGTTATATACATACATAGCATCAACTATTTT 857
QY 2569 CAAG-----AACCCCAATAGCAACAAACAAACAGACTAACAAAATGTGTAAACAAGA 2618
Db 858 CAAGAACAAATATTAAACCCGATAGCAACAAACAAACAGACTAACAAAATGTGTAAACAAGA 917
QY 2619 AACTAATGACCTTTCTAAATCAACATTCATATATCTACATGCTATTTTACAACAGG 2678
Db 918 AACTAATGACCTTTCTAAATCAACATTCATATATCTACATGCTTTTACAACAGG 977
QY 2679 GAAACCTCCATGGTTACAGGCATGTCATATGAAAATAAGCTGCAATAGC-TTTTAT 2737
Db 978 GAAACCTCCATGGTTACAGGCATGTCATATGAAAATAAGCTGCAATAGCTCAATTTAT 1037
QY 2738 ACAATTATCGCTCTCAAGAAAATGAATCATTAAGACAGTAATTAGGAGTTTCAAAATTTA 2797
Db 1038 ACAATTATCGCTCTCAAGAAAATGAATCATTAAGACAGTAATTAGGAGTTTCAAAATTTA 1097
QY 2798 AACATTTTACAGTAATTTTAAATATTGCTTCAATAATTTTAAATTTTGAAGTCTGAG 2857
Db 1098 AACATTTTACAGTAATTTTAAATATTGCTTCAATAATTTTGAAGTCTGAG 1145
QY 2858 TTTCAAAAGTGA--TTTTCCTCCCAAAAGTGCCCAACTTAAGCTAGAGCTTTCAAGTG 2914
Db 1146 TTTCAAAAGTGATTTTTCCTCCCAAAAGTTTCAACACTTAAGCTAGAGCTTTCAAGTG 1205
QY 2915 TTTACTTTGCCCT-AAAAGTTAAGACATATCTGAGAATCAATAAGTACACATGATTCT 2973
Db 1206 TTTACTTTGCCCTAAAAGTTAAGACATATCTGAGAATCAATAAGTACACATGATTCT 1265
QY 2974 GATGCTATCTGCTCTGTTTAATAACAAGATTCA 3007
Db 1266 GATGCTATCTGCTCTGTTTAATAACAAGTCTTTA 1299

RESULT 7
AK097915/c 2209 bp mRNA linear PRI 15-JUL-2002
LOCUS Homo sapiens cDNA FLJ40596 fis, clone THYM02010831, highly similar
DEFINITION to Mouse NCBP-29 mRNA for PW29.
ACCESSION AK097915
VERSION AK097915.1 GI:21757818
KEYWORDS oligo capping; fls (full insert sequence).
SOURCE Homo sapiens shymus cDNA to mRNA, clone_lib:THYM02
clone:THYM02010831.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Nishi,T., Ota,T., Nakagawa,S., Senoh,A., Mizuguchi,H., Inagaki,H.,
Sugiyama,T., Irie,K., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,
Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K.,
Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
TITLE Unpublished
JOURNAL
REFERENCE 2 (bases 1 to 2209)
AUTHORS Isogai,T. and Yamamoto,J.
DIRECT SUBMISSION
TITLE Direct Submission
JOURNAL Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team): 2-6-7
kazusa-kamatairi, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and

```

RAB: annotation: HRI and RAB.  
 FEATURES source 1..2209 /organism="Homo sapiens"  
 Location/Qualifiers /db\_xref="taxon:9606"  
 /clone="THYM2010831"  
 /tissue\_type="thymus"  
 /clone\_lib="THYM2"  
 /note="cloning vector: pME18SFL3"  
 BASE COUNT 678 a 387 c 440 g 704 t  
 ORIGIN

Query Match 11.4%; Score 356.2; DB 9; Length 2209;  
 Best Local Similarity 88.3%; Pred. No. 7.4e-50;  
 Matches 453; Conservative 0; Mismatches 33; Indels 27; Gaps 5;

QY 2510 ACAGCTTATACACACTTTTATTAGAAAGTTATACATAACACAGCATCAACTATTTC 2569  
 Db 2209 ACAGCTTATACACACTTTTATTAGAAAGTTATACATAACACAGCATCAACTATTTC 2150

QY 2570 AAG-----AACCCATAAGCAACAAACACAGACTAACAAATGTGAACAAGAA 2619  
 Db 2149 AAGCAACATATTAAACCGTAGCAACAAACACAGACTAACAAATGTGAACAAGAA 2090

QY 2620 ACTAATGACCTTTCTAAATCAACACTTCAATATATACAAATGCTCTATTACAAACAGGG 2679  
 Db 2089 ACTAATGACCTTTCTAAATCAACACTTCAATATATACAAATGCTCTTTTACAAACGGG 2030

QY 2680 AAACCTCCATGTTTACAGCATGTCATATTGAAATAAGCTGCAATAGC-TTTTATA 2738  
 Db 2029 AAACCTCCATGTTTACAGCATGTCATATTGAAATAAGCTGCAATAGCATTATA 1970

QY 2739 CAATTATCGCTCTCAAGAAATGAATCATTAGACAGTAATTAGGAGTTTCACAAATTTAA 2798  
 Db 1969 CAATTACCCTCTGAGAACTGAATCATTAACACAGTAATTACGAGTTTCACAAATTTAA 1910

QY 2799 ACATTTACAGTAATTTTAAATATTGCTCTCAATAATTTTAAATATTGAAGTCTGAGT 2858  
 Db 1909 ACATTTACATAATTTTAAATATTGCGT-----ATACACTGAAGTCTGAGT 1862

QY 2859 TTCAAAAGTGA---TTTTTCCACAAAGGTGCCACACTTAAGCTAGAGCTTTTCAGTGT 2915  
 Db 1861 TTCAAAAGTGAATTTTTCACAAAGGTGCCACACTTAAGCTAGAGCTTTTCAGTGT 1802

QY 2916 TAACCTTTGGCCT-AAAAGTTAAGACATATCTCGAGATCATATAATAGTCAATGTTCTGTG 2974  
 Db 1801 TAACCTTTGGCCTAAAAGTTAAGACATATCTCGATATCTGATATCATACAGTCAATGTTCTG 1742

QY 2975 ATGCTATCGCTCTGTTTAATAACAAGATTCA 3007  
 Db 1741 ATGCTATCGCTCTGTTTAATAATAAAGTCTTTA 1709

RESULT 8  
 HUMORF005/c  
 LOCUS Human mRNA for KIAA0078 gene, complete cds.  
 DEFINITION  
 ACCESSION D38551  
 VERSION D38551.1 GI:1531549  
 KEYWORDS KIAA0078.  
 SOURCE Homo sapiens male myeloblast cell-line KG-1 cDNA to mRNA.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1  
 AUTHORS Nomura.N., Nagase.T., Miyajima.N., Sazuka.T., Tanaka.A., Sato.S.,  
 Seki.N., Kawarabayashi.Y., Ishikawa.K. and Tabata.S.  
 TITLE Prediction of the coding sequences of unidentified human genes. II.  
 The coding sequences of 40 new genes (KIAA0041-KIAA0080) deduced by  
 analysis of cDNA clones from human cell line KG-1  
 JOURNAL DNA Res. 1 (5), 223-229 (1994)  
 MEDLINE 96051398  
 REFERENCE 2 (bases 1 to 3647)

AUTHORS Ohara,O., Nagase,T., Kikuno,R. and Nomura,N.  
 TITLE Direct Submission  
 JOURNAL Submitted (17-OCT-1994) Osamu Ohara, Kazusa DNA Research Institute;  
 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan  
 (E-mail:cdnainfo@kazusa.or.jp. Tel:++81-438-52-3913)  
 COMMENT On Sep 10, 1996 this sequence version replaced gi:559710.  
 FEATURES Location/Qualifiers  
 source 1..3647  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /sex="male"  
 /cell\_line="KG-1"  
 /cell\_type="myeloblast"  
 1..3647  
 /gene="KIAA0078"  
 1..184  
 /gene="KIAA0078"  
 185..2080  
 /gene="KIAA0078"  
 /note="The hal237 gene product is related to S.pombe rad21  
 gene product."  
 /codon\_start=1  
 /protein\_id="BAA07554.1"  
 /db\_xref="GI:1531550"  
 /translation="MFYHPLVSLKRGPLAKIWLAAHWDKLLTKAHVFECLNLSVESI  
 ISPKVMALRTSGHLLGVVRIYHRKAKYLLADCNFAFIKIKAFRPGVVDDLPENRE  
 AAYNAITLPEEFHDFDQDIDVAOQFSLNSRVEEITMREEVGNISILQENDF  
 GDFGMDREIMREGSAFEDDMVSTTPTSNLLSEQSTNLNEKINHLEYEQYKDD  
 SFGNDGGILDDKLIISNDNGIFDDPDLSEAGVLMPEQAPDDMDDEDVSMGGPD  
 SPQSDVPEMPTMDQTLVLPNEEFALEPIDITVKETKAKRKLIVDSVKEILDS  
 KTRIADLSYSDIVTTLDLAPPTKKLMHKEGTGVEKLFSLPAQPLNNRLLLFLTRC  
 LTPLVPEDLKRKRGGEADNLDFLEKFEVPEVREDOQQOHOQRVIDEPIILEPSR  
 LQESVMEASRTNIDESAMPPPPGQVKKRAGQIDPEVMPPOQVQVEQMEIPELPEE  
 PPNTCOLIPELELPEKEKEKEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE  
 LQALAKTAGAESISLLELCRNTNRKQAAKFYSFLVLKQQAIELTQEEPYSIIATP  
 GPRFHII"  
 2081..3647  
 /gene="KIAA0078"  
 1155 a 669 c 800 g 1023 t  
 BASE COUNT  
 ORIGIN

Query Match 11.4%; Score 356.2; DB 9; Length 3647;  
 Best Local Similarity 88.3%; Pred. No. 6.8e-50;  
 Matches 453; Conservative 0; Mismatches 33; Indels 27; Gaps 5;

QY 2510 ACAGCTTATACACACTTTTATTAGAAAGTTATACATAACACAGCATCAACTATTTC 2569  
 Db 3647 ACAGCTTATACACACTTTTATTAGAAAGTTATACATAACACAGCATCAACTATTTC 3588

QY 2570 AAG-----AACCCATAAGCAACAAACACAGACTAACAAATGTGAACAAGAA 2619  
 Db 3587 AAGCAACATATTAAACCGTAGCAACAAACACAGACTAACAAATGTGAACAAGAA 3528

QY 2620 ACTAATGACCTTTCTAAATCAACACTTCAATATATCTACAATGCTCTATTACAAACAGGG 2679  
 Db 3527 ACTAATGACCTTTCTAAATCAACACTTCAATATATCTACAATGCTCTTTTACAAACGGG 3468

QY 2680 AAAACTCCATGTTTACAGCATGTCATATTGAAAAATAAGCTGCAATAGC-TTTTATA 2738  
 Db 3467 AAAACTCCATGTTTACAGCATGTCATATTGAAAAATAAGCTGCAATAGCATTATA 3408

QY 2739 CAATTATCGCTCTCAAGAAATGAATCATTAAGACAGTAATTAGGAGTTTCACAAATTTAA 2798  
 Db 3407 CAATTACACACTCTCAAGAACTGAATCATTTAAACAGTAATTAGGAGTTTCACAAATTTAA 3348

QY 2799 AACATTTACAGTAATTTTAAATATTGCTCTCAATAATTTTAAATATTGAAAGTCTGAGT 2858  
 Db 3347 AACATTTACAGTAATTTTAAATATTGCGT-----ATACACTGAAGTCTGAGT 3300

QY 2859 TTCAAAAGTGA---TTTTTCCCAACAAGGTGCCAACACTTTAAGCTAGAGCTTTTCAGTGT 2915  
 Db 3299 TTCAAAAGTGAATTTTTCCTCCCAACAAGTTTCAACACTTTAAGCTAGAGCTTTTCAGTGT 3240



||||| ATGCTATCTGCTGCTGCTTAATAAAGTCTTTA 22472

# RESULT 10

AC087350/c

LOCUS

DEFINITION

AC087350

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

AC087350 193364 bp DNA linear PRI 21-AUG-2002  
Homo sapiens chromosome 8, clone RP11-367C15, complete sequence.  
AC087350  
HTG  
GI:22267854

human.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 193364)  
Birren,B., Nusbaum,C. and Lander,E.  
Homo sapiens chromosome 8, clone RP11-367C15  
Unpublished  
2 (bases 1 to 193364)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,  
Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B., Brown,A.,  
Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,  
Collamore,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S.,  
Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J.,  
Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,  
Jones,C., Karatas,A., Larocque,K., Lamazares,R., Lander,E.,  
Lehoczky,J., Levine,R., Liu,G., MacLean,C., Macdonald,P.,  
Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,  
McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,  
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,  
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,  
Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,  
Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,  
Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P.,  
Sounez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,  
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,  
Zembek,L., Zimmer,A. and Zody,M.

Direct Submission  
Submitted (28-DEC-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 193364)

Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,  
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,  
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,  
Cooke,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,  
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,  
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kanat,A.,  
Karatas,A., Kells,C., Lander,E., Levine,R., Lindblad-Toh,K.,  
Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,  
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,  
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,  
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,  
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,  
Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,  
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,  
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,  
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,  
Zembek,L., Zimmer,A. and Zody,M.

Direct Submission  
Submitted (16-AUG-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
4 (bases 1 to 193364)

Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,  
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,  
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,  
Cooke,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,  
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,

Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kanat,A.,  
Karatas,A., Kells,C., Lander,E., Levine,R., Lindblad-Toh,K.,  
Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,  
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,  
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,  
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,  
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,  
Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,  
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,  
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,  
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,  
Zembek,L., Zimmer,A. and Zody,M.

Direct Submission  
Submitted (21-AUG-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Aug 16, 2002 this sequence version replaced gi:16905274.  
All repeats were identified using RepeatMasker:  
Smith, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIDR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence.submissions@genome.wi.mit.edu  
Project Information  
Center project name: L11698  
Center clone name: 367\_C\_15

FEATURES  
source  
1. 193364  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8"  
/clone="RP11-367C15"  
/clone\_lib="RP11-11 Human Male BAC"  
repeat\_region  
290..348  
/rpt\_family="MIR"  
repeat\_region  
349..651  
/rpt\_family="AluSg"  
repeat\_region  
652..730  
/rpt\_family="MIR"  
repeat\_region  
1721..2290  
/rpt\_family="L1MB5"  
repeat\_region  
3577..3711  
/rpt\_family="MSTd"  
repeat\_region  
3728..3786  
/rpt\_family="FRAM"  
repeat\_region  
3812..4016  
/rpt\_family="MSTd"  
repeat\_region  
complement(4135..4438)  
/rpt\_family="AluY"  
repeat\_region  
4531..4597  
/rpt\_family="(TTCC)n"  
repeat\_region  
5406..6437  
/rpt\_family="L1MEC"  
repeat\_region  
6438..6744  
/rpt\_family="AluY"  
repeat\_region  
6745..7328  
/rpt\_family="L1MEC"  
repeat\_region  
8016..8321  
/rpt\_family="AluSx"  
repeat\_region  
complement(8338..8431)  
/rpt\_family="L3"  
repeat\_region  
complement(9625..9835)  
/rpt\_family="MLT1C"  
repeat\_region  
complement(9836..10138)  
/rpt\_family="AluX"  
repeat\_region  
complement(10139..10431)  
/rpt\_family="MLT1C"  
repeat\_region  
10441..11704  
/rpt\_family="L2"  
repeat\_region  
12286..12710

```

/rpt_family="LIPAL6"
12715..12735
/rpt_family="TAAA)n"
12806..12829
/rpt_family="HAL1"
12830..13136
/rpt_family="AluSx"
13137..13265
/rpt_family="HAL1"
13791..13843
/rpt_family="L2"
complement(14375..14530)
/rpt_family="MIR3"
15285..16638
/rpt_family="PTR5"
complement(16805..17197)
/rpt_family="MLTIL"
complement(17353..17462)
/rpt_family="MIR"
complement(17890..18020)
/rpt_family="MIR3"
18233..18405
/rpt_family="MIR"
complement(18743..18834)
/rpt_family="MIR"
18857..18915
/rpt_family="MIR"
complement(21125..21435)
/rpt_family="AluSx"
21499..21528
/rpt_family="AT-rich"
complement(22565..22867)
/rpt_family="AluY"
23040..23060
/rpt_family="AT-rich"
24278..24530
/rpt_family="L2"
24532..24562
/rpt_family="(TTAA)n"
complement(24678..24746)
/rpt_family="Charlie4"
complement(25284..25429)
/rpt_family="MIR"
25489..25526
/rpt_family="5S"
complement(25586..25737)
/rpt_family="MIR3"
complement(25941..26235)
/rpt_family="AluSq"
27079..27100
/rpt_family="(CCG)n"
28055..28463
/rpt_family="MSTA"
28644..28671
/rpt_family="(TTCA)n"
complement(28724..28949)
/rpt_family="MIR"
complement(29453..29575)
/rpt_family="L2"
complement(29580..29674)

Query Match 11.1%; Score 346.6; DB 9; Length 193364;
Best Local Similarity 87.1%; Pred. No. 1.4e-48;
Matches 447; Conservative 0; Mismatches 39; Indels 27; Gaps 5;

QY 2509 TACAGCTTAAACACAACTTTATTAGAAAGTTATACATACACAGCATCAACTATTTT 2568
|||||
Db 119652 TACAGCTTAAACACAACTTTATTAGAAAGTTATACATACACAGCATCAACTATTTT 119593
|||||

QY 2569 CAAG-----AACCAATAAGCAACAAACACAGACTTAACAAAATGTTGAACAAGA 2618
|||||
Db 119592 CAAGAACATATTAAACCCGATAGCAACAAACACAGACTTAACAAAATGTTGAACAAGA 119533
|||||

```

QY 2619 AACTAATGACCTTTCTAAATAACAACTTCAATTTATCTACAAATGCTATTTTACAAACAGG 2678  
 |||||  
 Db 119532 AACTAATGACCTTTCTAAATAACAACTTCAATTTATCTACAAATGCTATTTTACAAACGGG 119473  
 |||||  
 QY 2679 GAAACTCCATGGTTTACAGGATGTCATATTTGAAAATAAGCTGCAATAGC-TTTTAT 2737  
 |||||  
 Db 119472 GAAACTCCATGGTTTACAGGATGTCATATTTGAAAATAAGCTGCAATAGCAATTTAT 119413  
 |||||  
 QY 2738 ACAATTATCGCTCTCAAGAAATGAATCATTAAGACATTAATTAGGAGTTTCACAAATTA 2797  
 |||||  
 Db 119412 ACAATTATCGCTCTCAAGAAATGAATCATTAAGACATTAATTAGGAGTTTCACAAATTA 119353  
 |||||  
 QY 2798 AAACATTTCCAGTAAATTTTAAATATTTGCTTCAATAATTTTAAATATTTGAAGTCTGAG 2857  
 |||||  
 Db 119352 AAACATTTCCAGTAAATTTTAAATATTTGCTTCAATAATTTTAAATATTTGAAGTCTGAG 119305  
 |||||  
 QY 2858 TTTCAAAAGTGA- -TTTTCCTCCCAAGAGTGCAACACTTAAGCTAGAGCTTTCAGTG 2914  
 |||||  
 Db 119304 TTTCAAAAGTGAATTTTTCCTCCCAAGAGTTTCAACACTTAAGCTAGAGCTTTCAGTG 119245  
 |||||  
 QY 2915 TTAACCTTTCCTCCCTAAAGTTAAGACATATTTCTGAGAATCATATAGTCACATGATTTCTG 2974  
 |||||  
 Db 119244 TTAACCTTTCCTCCCTAAAGTTAAGACATATTTCTGAGAATCATATAGTCACATGATTTCTG 119186  
 |||||  
 QY 2975 ATGCTATCTGCTCTGTTAATAACAAAGATTCA 3007  
 |||||  
 Db 119185 ATGCTATCTGCTCTGTTAATAACAAAGTTCTTTA 119153  
 |||||

RESULT 11  
 HSU85A3  
 LOCUS  
 DEFINITION  
 Human DNA sequence from clone U85A3 on chromosome X contains CSF2  
 (cleavage stimulation factor, subunit 2, 64kD) pseudogene similar  
 to T-cell cyclophilin, pseudogene similar to S.pombe rad21, STS  
 and GSSs, complete sequence.

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

278021  
 278021.1 GI:1483514  
 HTG; CSF2.  
 Homo sapiens.  
 Mammalia; Euthera; Primates; Catarrhini; Homidae; Homo.  
 1 (bases 1 to 37652)  
 Odell.C.  
 Direct Submission  
 Submitted (31-JUL-1996) Sanger Centre, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
 requests: clonerequest@sanger.ac.uk  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.  
 This sequence has been finished according to sequencing map criteria  
 as follows. An attempt is made to resolve all sequencing problems,  
 such as compressions and repeats, but not necessarily within known  
 annotated human repeat sequence elements (e.g. Alu). Where the  
 sequence is ambiguous, there is an annotation using the 'unsure'  
 feature key.

This sequence was generated from part of bacterial clone contigs of  
 human chromosome X, constructed by the Sanger Centre Chromosome X  
 Mapping Group. Further information can be found at  
 http://www.sanger.ac.uk/HGP/ChrX  
 U85A3 is from the Lawrence Livermore National Laboratory  
 flow-sorted X chromosome cosmid library LLOXNC01.

FEATURES  
 Source  
 1..37652  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="X"  
 /clone="LLOXNC01-85A3"

```

/cione_lib="LLOXNC01"
228..285
/note="Alu repeat: matches 247..304 of consensus"
/evidence-not_experimental
repeat_region
gene
complement(join(673..842,1488..1566,2619..2676))
/gene="CSTF2"
complement(join(<673..842,1488..1566,2619..2676))
/gene="CSTF2"
/note="match: M85085"
/evidence-not_experimental
repeat_region
complement(join(<673..842,1488..1566,2619..2676))
/gene="CSTF2"
/note="match: protein P33240"
/codon_start=1
/evidence-not_experimental
/product="CSTF2 (cleavage stimulation factor, subunit 2,
64kD)"
/protein_id="CA841262.1"
/db_xref="GI:4678893"
/translation="MAGLTVPDPAVDRSLRSVFGNIPYEATEQLKDIFSEVGPVVS
FRIVYDREGTKPGYGCYQDETALSAMRNLRNFGESGRALRVDNAASEKNKEELK
983..1281
/note="AluX repeat: matches 1..300 of consensus"
/evidence-not_experimental
repeat_region
1324..1369
/note="L1MB8 repeat: matches 6126..6173 of consensus"
/evidence-not_experimental
misc_feature
complement(1919..2597)
/gene="CSTF2"
/note="match: GSS B68800 clone 2023F10"
3581..3870
/note="AluSp repeat: matches 1..293 of consensus"
/evidence-not_experimental
repeat_region
4175..4354
/note="MER5A repeat: matches 1..178 of consensus"
/evidence-not_experimental
repeat_region
4787..5081
/note="MLT1E repeat: matches 264..568 of consensus"
/evidence-not_experimental
repeat_region
5208..5278
/note="MLT2FA repeat: matches 118..198 of consensus"
/evidence-not_experimental
repeat_region
5391..5513
/note="MIR repeat: matches 11..144 of consensus"
/evidence-not_experimental
repeat_region
7362..7921
/note="L1MB3 repeat: matches 5614..6182 of consensus"
/evidence-not_experimental
repeat_region
7922..8420
/note="MLT2CA repeat: matches 1..503 of consensus"
/evidence-not_experimental
repeat_region
8422..9916
/note="L1MB3 repeat: matches 4109..5617 of consensus"
/evidence-not_experimental
repeat_region
9914..9994
/note="L1M4c repeat: matches 1853..1933 of consensus"
/evidence-not_experimental
repeat_region
10004..10872
/note="L1PBA repeat: matches 1355..412 of consensus"
/evidence-not_experimental
repeat_region
10860..10971
/note="L1PBA repeat: matches 279..160 of consensus"
/evidence-not_experimental
repeat_region
10975..13231
/note="L1PBI repeat: matches 1058..3300 of consensus"
/evidence-not_experimental
repeat_region
13232..13536
/note="AluY repeat: matches 1..304 of consensus"
/evidence-not_experimental
repeat_region
13537..14462
/note="L1PBI repeat: matches 3299..4224 of consensus"
/evidence-not_experimental
14464..14947
/note="L1R10A repeat: matches 1..548 of consensus"
/evidence-not_experimental
repeat_region
14973..16855
/note="L1PBI repeat: matches 4234..6151 of consensus"
/evidence-not_experimental
repeat_region
16876..16998
/note="L1M4c repeat: matches 1712..1835 of consensus"
/evidence-not_experimental
repeat_region
17090..17924
/note="L1M4c repeat: matches 846..1603 of consensus"
/evidence-not_experimental
repeat_region
18176..18298
/note="L1M3d repeat: matches 495..618 of consensus"
/evidence-not_experimental
repeat_region
18299..18460
/note="L1M4 repeat: matches 79..236 of consensus"
/evidence-not_experimental
misc_feature
18501..18903
/note="match: STS G23519"
complement(19640..20032)
/note="match: STS G11279"
complement(20033..21918)
/gene="CU85A3.3"
complement(20033..21918)
/gene="CU85A3.3"
/note="match: D38551"
/pseudo
/codon_start=1
/evidence-not_experimental
/product="CU85A3.3 (similar to S.pombe rad21)"
20346..20787
/note="match: STS G30625"
20368..20616
/note="match: STS G29205"
22030..22270
/note="MER4C repeat: matches 143..447 of consensus"
/evidence-not_experimental
gene
22311..22799
/gene="CU85A3.2"
22311..22799
/gene="CU85A3.2"
/note="match: P05092
match: Y00052"
/pseudo
/codon_start=1
/evidence-not_experimental
/product="CU85A3.2 (similar to T-cell cyclophilin)"
23010..23320
/note="AluSp repeat: matches 1..311 of consensus"
/evidence-not_experimental
repeat_region
23375..23691
/note="MER4A repeat: matches 3..337 of consensus"
/evidence-not_experimental
repeat_region
23695..23867
/note="MER4-internal repeat: matches 6450..6598 of
consensus"
/evidence-not_experimental
repeat_region
23925..23982
/note="29 copies 2 mer aa 76% conserved"
/evidence-not_experimental
repeat_region
24322..24501
/note="HERV1L18 repeat: matches 5621..5801 of consensus"
/evidence-not_experimental
repeat_region
26573..26872
/note="AluSq repeat: matches 1..301 of consensus"
/evidence-not_experimental
repeat_region
27271..27604
/note="AluYa8 repeat: matches 1..307 of consensus"
/evidence-not_experimental
repeat_region
27646..27906
/note="HUERS-P3 repeat: matches 4101..5118 of consensus"
/evidence-not_experimental

```





```
variation
6020. .6022
/replace="clone U85A3; GAC in this entry; substitution"
/replace="ggc"
6435. .6584
/replace="LRR7 repeat; matches 448. .301 of consensus"
complement(6465. .11676)
/notes="remnants of endogenous retrovirus; partial homology
to viral RTVL-H, Phospholipase A2 homologue, GAG core shell
proteins, envelope (coat) proteins nucleic acid binding
proteins, POL polyprotein, reverse transcriptase,
ribonuclease H; matches many proteins, ESTs, cDNAs,
genomic DNAs and STSS"
7297. .7299
/notes="clone U85A3; AAA in this entry; substitution"
/replace="ata"
8284. .8286
/replace="ata"
/notes="clone U85A3; GCA in this entry; substitution"
/replace="gga"
8570. .8572
/replace="gga"
8622. .8624
/replace="gga"
/notes="clone U85A3; TAG in this entry; substitution"
/replace="tgg"
9488. .9490
/replace="tgg"
/notes="clone U85A3; AAT in this entry; substitution"
/replace="agt"
11025. .11027
/replace="agt"
/notes="clone U85A3; TCT in this entry; substitution"
/replace="ttt"
11321. .11620
/notes="LRR7 repeat; matches 448. .134 of consensus"
11572. .11681
/notes="LRR7 repeat; matches 131. .1 of consensus"
13573. .18746
/notes="remnants of endogenous retrovirus; partial homology
to viral envelope (coat) proteins, POL polyprotein,
reverse transcriptase, ribonuclease H; matches many
proteins, ESTs, cDNAs and genomic DNAs"
13669. .13671
/replace="gtt"
/notes="clone U85A3; GCT in this entry; substitution"
/replace="gtt"
13986. .13988
/replace="gtt"
/notes="clone U85A3; TAT in this entry; substitution"
/replace="tgt"
15110. .15112
/replace="tgt"
/notes="clone U85A3; CCG in this entry; substitution"
/replace="ctg"
15318. .15640
/notes="AluY repeat; matches 1. .299 of consensus"
16051. .16349
/notes="AluSg repeat; matches 2. .301 of consensus"
16632. .16634
/replace="acg"
/notes="clone U85A3; ATG in this entry; substitution"
/replace="acg"
17593. .17595
/replace="acg"
/notes="clone U85A3; AAA in this entry; substitution"
/replace="aga"
17984. .17986
/replace="aga"
/notes="clone U85A3; TTC in this entry; substitution"
/replace="tgc"
18981. .18982
/replace="tgc"
/notes="clone U85A3; CT in this entry; deletion"
/replace="ctt"
19001. .19117
/replace="ctt"
/notes="3 copies of 39 mer 82 % conserved"
19146. .19148
/replace="ctt"
/notes="clone U85A3; ATC in this entry; substitution"
/replace="agc"
19225. .19541
/replace="agc"
/notes="MER4A repeat; matches 3. .337 of consensus"
19439. .19441
/replace="agc"
/notes="clone U85A3; ATC in this entry; substitution"
```

```
repeat_region
variation
gene
mRNA
/replace="acc"
19800. .19902
/notes="AluSp repeat; matches 303. .1 of consensus"
19607. .19608
/replace="clone U85A3; AC in this entry; deletion"
/replace="atttac"
complement(19787. .20634)
/replace="dJ347M6.1"
/genes="dJ347M6.1"
complement(<19787. .>20619)
/genes="dJ347M6.1"
/notes="match: cDNAs M25637 M19533 Y00052 X52803 X17105;
match: ESTs AA075585 AA299048 F04720 AA362886 AA488379
AA486598 AA311324 AA131423 T70141 AA307700 AA533008
AA513009 AA180415 AA314272 AA081118 AA300395 AA505262
D57979 AA353507 AA581820 AA126791 AA401598 AA335880
AA573841 AA171787 AA582439 AA299336 D55800 HA3341 F18530
D55283 AA315532 AA504064 AA182044 AA191632 AA302188
AA342661 AA191634 AA220938 AA126843 AA532728 AA488511
AA088390 AA307477 AA412697 AA576164 AA342786 AA147099
AA155677 AA552685 AA466693 R87887 W79891 AA101162 AA314411
AA321796 AA483724 AA373538 AA356609 AA147872 AA531206
AA070523 AA305622 AA307522 AA554115 AA346116 T69131
AA148353 AA522575 AA306872 AA308538 W87520 AA173745
AA398530 AA488404 W87521 AA157745 AA528078 AA385581
AA146938 AA313332 AA302560 AA328140 AA191136 AA306921
AA370383 AA181295 AA328545 AA321796 AA483724 AA373538
AA356609 AA147872"
/pseudo
/evidence=not_experimental
2009. .20011
/notes="clone U85A3; AGA in this entry; substitution"
/replace="aca"
complement(20116. .20634)
/genes="dJ347M6.1"
Query Match 10.2%; Score 317.6; DB 9; Length 56583;
Best Local Similarity 83.3%; Pred. No. 1.le 43;
Matches 418; Conservative 0; Mismatches 59; Indels 25; Gaps 4;
Qy 2504 TTTTTCACAGCTTATACACAACTTTTATTAGAAAGTTATACATAACACAGCATCAACT 2563
Db 24425 TTGTTTCAACTATATACACAACTTTTATTAGAGAGTTATACATAACATGATCAACT 24366
Qy 2564 ATTTTCAGAAC-----CCAAATAAGCAACAAAACACAGACTAACAAATGTGTAA 2614
Db 24365 ATTTTCAAGAACAAATATTAACCCCAATTAAGCAAAAAACACAGACTAACAAATGTGTAA 24306
Qy 2615 AAGAACTTAATGACCTTCTTAAATCAACATTCGAATATCTACAATGCTCTATTACAAA 2674
Db 24305 AAGAAACCAATGACCTTCTTAAATCAACATTCGAATATCTACAATGCTCTATTACAAA 24246
Qy 2675 CAGGAAAACCTCCATGGTTTACAGGCATGTCATATTGAAAATAAAGCTGCAATAGCTTTT 2734
Db 24245 CAGGAAAACCTCCATGGTTTACAGGCACAGCATATTGAAATAAATCTGCAATAGCAAT 24186
Qy 2735 TATACAATTTTCGCTCTCAAGAAATGAATCATTAAGACAGTAATAGGAGTTCACAAAT 2794
Db 24185 TTTTATGGTTAAACCTCTCAGAAACGGAATCATTTAAACAGTAAT---GAGTTTCACAAAT 24129
Qy 2795 TTAACAACTTTTCAGTAAATTTTAAATTTATTCCTCAATATTTTAAATTTTAAAGTCT 2854
Db 24128 TTAACAACTTTTCAGTAAATTTTAAATTTATTCAGT-----ATACACTGAAGTCT 24081
Qy 2855 GAGTTTCAAAAGTGATTTTTTCCCAAAAGTGCCAACTTAAAGCTAGAGCTTTTCAGTG 2914
Db 24080 GTTTCAAAAGTGATTTTTTCCCAAAAGTGCCAACTTAAAGCTAGAGCTTTTCAGTG 24021
Qy 2915 TTAACCTTGCCCT-AAAAGTTAAGACATATTTCTGAGATCATATAATAGTCACATGATTTCT 2973
Db 24020 TTAATTTTGCCCTAAAAAGTTAAGACATATTTCTGATATATAATAGTCACATGATTTCT 23961
```





```
Db 544 CCAATAACCTTGAAGAGCGTTGCGGATAATAAATTCGGTTCTTTCAGAAGAATGGGCT 603
QY 719 GGAATGATCTTATCTGTGAACTAGAGGAATTCATATGTGAGATGAATAAGATTACC 778
Db 604 GGAATGACATTCAGTGTATGTACCTCAGAGTCAATTTGCAAGATGAAGAAGATCTACA 663
QY 779 TATGAGTAGAAGCTT 793
Db 664 TATAATGAAATATT 678
```

Search completed: February 17, 2003, 11:03:43  
Job time : 6410.63 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 17, 2003, 08:48:06 ; Search time 16.425 seconds  
(without alignments)  
11706.916 Million cell updates/sec

Title: US-09-766-511b-52  
Perfect score: 627  
Sequence: 1 atgatgaagagcagcaacc.....agatgaataagatttaacctta 627

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA: \*  
1: /cgn2\_6/ptodata/1/ina/5A-COMB.seq: \*  
2: /cgn2\_6/ptodata/1/ina/5B-COMB.seq: \*  
3: /cgn2\_6/ptodata/1/ina/6A-COMB.seq: \*  
4: /cgn2\_6/ptodata/1/ina/6B-COMB.seq: \*  
5: /cgn2\_6/ptodata/1/ina/PTUS-COMB.seq: \*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	373.2	59.5	1227	3	US-08-772-440-3
2	303.2	48.4	501	3	US-08-772-440-20
3	260.2	41.5	393	3	US-08-772-440-22
4	192	30.6	1104	4	US-09-111-470-1
5	134	21.4	1418	4	US-09-111-470-7
6	94.6	15.1	145	3	US-08-772-440-42
7	94.4	15.1	152	3	US-08-772-440-40
8	94.4	15.1	10409	3	US-08-772-440-33
9	79.2	12.6	116	3	US-08-772-440-41
10	60	9.6	101	3	US-08-772-440-39
11	58.4	9.3	92	3	US-08-772-440-38
12	58	9.3	84	3	US-08-772-440-18
13	52	8.3	10409	3	US-08-772-440-33
14	44.2	7.0	871	1	US-08-650-578-1
15	40.2	6.4	301	4	US-09-222-575-29
16	36	5.7	1370	4	US-09-111-470-9
17	36	5.7	1458	4	US-09-111-470-3
18	35.4	5.6	360	4	US-08-543-246B-4
19	35.4	5.6	405	4	US-08-543-246B-3
20	35.4	5.6	645	4	US-08-543-246B-12
21	35.4	5.6	699	4	US-08-543-246B-11
22	35.4	5.6	1333	4	US-08-543-246B-15
23	35.4	5.6	1387	4	US-08-543-246B-1
24	33	5.3	1615	1	US-09-640-305-5
25	33	5.3	1615	1	US-08-360-673-5
26	32.8	5.2	405	4	US-08-543-246B-7
27	32.8	5.2	693	4	US-08-543-246B-13

c	28	32.8	5.2	1222	4	US-08-543-246B-5	Sequence 5, Appli
	29	31.8	5.1	6349	4	US-09-484-970B-114	Sequence 114, App
	30	31.4	5.0	874	3	US-08-462-436-1	Sequence 1, Appli
	31	31.4	5.0	874	3	US-08-465-275-1	Sequence 1, Appli
	32	31.4	5.0	874	4	US-08-840-877-1	Sequence 7, Appli
	33	31.4	5.0	1213	4	US-08-976-259-7	Sequence 68, Appli
c	34	31.4	5.0	3241	4	US-08-232-463-14	Sequence 14, Appl
c	35	31.4	5.0	7218	1	US-09-641-638-49	Sequence 49, Appl
	36	31	4.9	448	4	US-09-641-638-50	Sequence 50, Appl
	37	31	4.9	448	4	US-09-641-638-51	Sequence 51, Appl
	38	31	4.9	448	4	US-09-641-638-571	Sequence 571, App
	39	30.8	4.9	4308	3	US-08-938-291A-8	Sequence 8, Appli
	40	30.6	4.9	1789	1	US-08-393-219-1	Sequence 1, Appli
	41	30.4	4.8	1025	1	US-08-365-103B-9	Sequence 9, Appli
	42	30.4	4.8	1037	1	US-08-365-103B-7	Sequence 7, Appli
	43	30.4	4.8	1212	4	US-09-591-435-10	Sequence 10, Appl
c	44	30.4	4.8	3846	2	US-08-845-161A-5	Sequence 5, Appli
	45	30.4	4.8	3846	2	US-08-845-161A-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1  
US-08-772-440-3  
; Sequence 3, Application US/08772440  
; Patent No. 6046158  
; GENERAL INFORMATION:  
; APPLICANT: Arizumi, Kiyoshi  
; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE  
; TITLE OF INVENTION: LECTINS, DEXTIN-1 AND DEXTIN-2; COMPOSITIONS AND USES  
; TITLE OF INVENTION: THEREOF  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/772.440  
; FILING DATE: CONCURRENTLY HEREWITH  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Parker, David L.  
; REGISTRATION NUMBER: 32,165  
; REFERENCE/DOCKET NUMBER: UTXD:493  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 512/418-3000  
; TELEFAX: 512/474-7577  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1227 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-772-440-3

Query Match 59.5%; Score 373.2; DB 3; Length 1227;  
Best Local Similarity 76.1%; Pred. No. 2.4e-109;  
Matches 475; Conservative 0; Mismatches 143; Indels 6; Gaps 1;  
QY 10 GAGCAGCAACCTCAACAGATACAGAGAAAGAGCTGGTGTCCCTGACACTCTGGTCTGTG 69  
Db 149 GTGCAGGAAGACATCCCAAGGAGGAGTCTGCTGGACCCCTGACACTCTGGTCTGTG 208

```

1  RESULT 2
2  US-08-772-440-20
3  ; Sequence 20, Application US/08772440
4  ; Patent No. 6046158
5  ; GENERAL INFORMATION:
6  ; APPLICANT: Ariizumi, Kiyoshi
7  ; APPLICANT: Takashima, Akira
8  ; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE
9  ; TITLE OF INVENTION: LECTINS, DEXTIN-1 AND DEXTIN-2; COMPOSITIONS AND USES
10 ; TITLE OF INVENTION: THEREOF
11 ; NUMBER OF SEQUENCES: 42
12 ; CORRESPONDENCE ADDRESS:
13 ; ADDRESSEE: Arnold, White & Durkee
14 ; STREET: P.O. Box 4433
15 ; CITY: Houston
16 ; STATE: Texas
17 ; COUNTRY: USA
18 ; ZIP: 77210
19 ; COMPUTER READABLE FORM:
20 ; MEDIUM TYPE: Floppy disk
21 ; COMPUTER: IBM PC compatible
22 ; OPERATING SYSTEM: PC-DOS/MS-DOS
23 ; SOFTWARE: Patentin Release #1.0, Version #1.30
24 ; CURRENT APPLICATION DATA:
25 ; APPLICATION NUMBER: US/08/772,440
26 ; FILING DATE: CONCURRENTLY HERewith
27 ; CLASSIFICATION: 435
28 ; ATTORNEY/AGENT INFORMATION:
29 ; NAME: Parker, David L.
30 ; REGISTRATION NUMBER: 32,165
31 ; REFERENCE/DOCKET NUMBER: UTXD:493
32 ; TELECOMMUNICATION INFORMATION:

```

```

RESULT 3
US-08-772-440-22
: Sequence 22, Application US/08772440
: Patent No. 6046158
: GENERAL INFORMATION:
: APPLICANT: Ariizumi, Kiyoshi
: APPLICANT: Takashima, Akira
: TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE
: TITLE OF INVENTION: LECTINS, DECTIN-1 AND DECTIN-2; COMPOSITIONS AND USES
: TITLE OF INVENTION: THEREOF
: NUMBER OF SEQUENCES: 42
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Arnold, White & Durkee
: STREET: P.O. Box 4433
: CITY: Houston
: STATE: Texas
: COUNTRY: USA
: ZIP: 77210
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/772,440
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTXD:493
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-772-440-22

Query Match          41.5%; Score 260.2; DB 3; Length 393;
Best Local Similarity 78.9%; Pred. No. 1.6e-73;
Matches 310; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 235 TGCCAGCTTCTTGGAAGTCATTTGGTCCAGTTGCTTACTTTCATTTCCAGTGAAGAGAAG 294
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 TGCCAAATCACTGGAAGTCATTTGGCTCCAGCTGCTACCTCATTTCTACCAAGGAGAAC 60

QY 295 GTTTGGTCTAAGAGTACGAGAACTGTGTTGAGATGGGAGCACATTTGGTTGTTCAAC 354
      || | | | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 TTCTGGAGCACGACGTAGCAGAACTGTGTTTCAGATGGGGCTCATCTGTTGGTGATCAAT 120

QY 355 ACAGAAGCAGACGAGAAATTTTCATTTGCCAGCAGCTGAATCAGTCATTTTCTTATTTCTG 414
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 ACTGAAGCGGAGCAGAAATTTTCATCACCCAGCAGCTGAATCAGTCATTTTCTTACTTCCTG 180

QY 415 GGGCTTTTCAGACCCACAAAGTAATAATTAATTTGGCAATGGATGATGAAGACACCTTATGAG 474
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 GGTCTTTCCGATCCACAAAGTAATGGCAATGGCAATGGATGATGATCTCTCTTCAGT 240

QY 475 AAAATGTCAGATTTTGGCACCTAGGTGAGGCCCAATCAATTTCTCGACAGCAATGTGCTTCA 534
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 CAAAATGTCAGGTCTGGCACCCCATGAAACCAATCTTCCAGAAAGCGGTGTGTTTCA 300

QY 535 ATAGTCTTCTGGAACCTACAGGATGGGCTGGAATGATGTTATCTGTGAACTAGAAGG 594
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 ATAGTTTACTGGATCTCTCGAATGGGCTGGAATGATGTTTCTCTGTAGTAGTAACAC 360

QY 595 AATTCAATATGTGAGATGAATAAGATTACCTA 627
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 AATTCAATATGTGAATGAAGAAGATTACCTA 393
```

```

RESULT 4
US-09-111-470-1
; Sequence 1, Application US/09111470
; Patent No. 6277959
; GENERAL INFORMATION:
; APPLICANT: Valladeau, Jenny
; APPLICANT: Ravel, Odile
; APPLICANT: Bates, Elizabeth E.M.
; APPLICANT: Ford, John
; APPLICANT: Saeland, Sem
; APPLICANT: Lebecque, Serge J.E.
; TITLE OF INVENTION: Mammalian Membrane Protein Genes;
; TITLE OF INVENTION: Related Reagents
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESS: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
```

```

; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/111,470
; FILING DATE: 08-JUL-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/053,080
; FILING DATE: 09-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: SF0695
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)852-9196
; TELEFAX: (650)496-1200
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1104 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 242..952
; US-09-111-470-1

Query Match          30.6%; Score 192; DB 4; Length 1104;
Best Local Similarity 68.0%; Pred. No. 1.5e-51;
Matches 283; Conservative 0; Mismatches 137; Indels 3; Gaps 1;

QY 215 AGTGCCAGCCCTGGGGATGTTGCCAGCTTCTTGAAGTCATTTGGTTCACGTTGCTACT 274
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 537 AGAGACACCCCTGGAGCTGTTGCCCAAGAATTTGGAAGTCATTTAGTTCACACTGCTACT 596

QY 275 TCATTTCCAGTGAAGAGAAGGTTTGGTCTAAGAGTGAGCAGAACTGTGTTGAGATGGGAG 334
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 597 TTATTTCTACTGAATCAGCATCTTGCAAGACAGTGAGAAGGACTGTGCTAGAATGGAGG 656

QY 335 CACATTTGGTCTGTCTCAACAGAGCAGAGCAATTTTCATTTCCAGCAGCTGAATG 394
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 657 CTCACCTGCTGTGATGAACAACACTCAAGAAGAGCAGGATTTCACTTCCAGAATCTGCAAG 716

QY 395 AGTCATTTTCTTATTTTCTGGGGCTTTCAGACCCACAAAGTAATAATAATTGGCAATGGA 454
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 717 AAGAATCTGCTTATTTTGTGGGCTCTCAGATCCAGAGGTCAGCGACATTTGGCAATGGG 776

QY 455 TTGATAAGACACCTTATGAGAAAAATGTCAGATTTTGGCACCTAGGTGAGCCCAATCATTT 514
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 777 TTGATCAGACACCACATACAAATGAAGTTCCACATTTCTGGCATCCAGGTGAGCCAGTGATC 836

QY 515 CTGCAGAGCAATGCTCTCAATAGCTTCTGCAA--ACCTACAGSATGGGCTGGAATG 571
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 837 CCAATGAGCGCTGCGTTGTGCTAAATTTTCGTAATATCCCAAAAAGATGGGCTGGAATG 896

QY 572 ATGTTATCTGTGAACACTAGAAGGAATTCATATGTGAGATGAATAAGATTTACCTA 627
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 897 ATGTTAATTTGCTTGTGCTCTCAAGGTCAGTTTGTGAGATGATGAAGATCCACTTA 952

RESULT 5
US-09-111-470-7
; Sequence 7, Application US/09111470
; Patent No. 6277959
; GENERAL INFORMATION:
; APPLICANT: Valladeau, Jenny
; APPLICANT: Ravel, Odile
; APPLICANT: Bates, Elizabeth E.M.
```

```

> APPLICANT: Ford, John
> APPLICANT: Saeland, Sem
> APPLICANT: Lebecque, Serge J.E.
> TITLE OF INVENTION: Mammalian Membrane Protein Genes;
> TITLE OF INVENTION: Related Reagents
> NUMBER OF SEQUENCES: 11
> CORRESPONDENCE ADDRESS:
> ADDRESSEE: DNAX Research Institute
> STREET: 901 California Avenue
> CITY: Palo Alto
> STATE: California
> COUNTRY: USA
> ZIP: 94304-1104
> COMPUTER READABLE FORM:
> MEDIUM TYPE: Floppy disk
> COMPUTER: IBM PC compatible
> OPERATING SYSTEM: PC-DOS/MS-DOS
> SOFTWARE: Patent In Release #1.0, Version #1.30
> CURRENT APPLICATION DATA:
> APPLICATION NUMBER: US/09/111,470
> FILING DATE: 08-JUL-1998
> CLASSIFICATION:
> PRIOR APPLICATION DATA:
> APPLICATION NUMBER: US 60/053,080
> FILING DATE: 09-JUL-1997
> ATTORNEY/AGENT INFORMATION:
> NAME: Ching, Edwin P.
> REGISTRATION NUMBER: 34,090
> REFERENCE/DOCKET NUMBER: SF0695
> TELECOMMUNICATION INFORMATION:
> TELEPHONE: (650)852-9196
> TELEFAX: (650)496-1200
> INFORMATION FOR SEQ ID NO: 7:
> SEQUENCE CHARACTERISTICS:
> LENGTH: 1418 base pairs
> TYPE: nucleic acid
> STRANDEDNESS: single
> TOPOLOGY: linear
> MOLECULE TYPE: cDNA
> FEATURE:
> NAME/KEY: CDS
> LOCATION: 279..992
> FEATURE:
> NAME/KEY: misc_feature
> LOCATION: 1348
> OTHER INFORMATION: /note= "poly-A addition motif"
> PS-09-111-470-7

```

[illegible]

```

Db 881 CAATGAAATAATGCTACATAAATTTACCGTTGGAAGACTGGATGGGGCTGGAACGATAT 940
Qy 576 TATCTGTGAAACTAGAGGAATTCATATATGTGAGATGAATAAGATTTACCTTA 627
Db 941 CTCITGCAGTCTTAAACAGAAAGTCAGTTTGTGCAGATGAAGAAAAATAAACTTA 992

RESULT 6
US-08-772-440-42
; Sequence 42, Application US/08772440
; Patent No. 6046158
; GENERAL INFORMATION:
; APPLICANT: Ariizumi, Kiyoshi
; APPLICANT: Takashima, Akira
; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE
; TITLE OF INVENTION: LECTINS, DECTIN-1 AND DECTIN-2; COMPOSITIONS AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/772,440
; FILING DATE: CONCURRENTLY HEREWITH
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTXD:493
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-772-440-42

Query Match 15.1%; Score 94.6; DB 3; Length 145;
Best Local Similarity 79.4%; Pred No. 6.7e-21;
Matches 112; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 487 TTTTGGCACCTAGGTAGCCCAATCATCTCGACGAGCAATGTGCTTCAATAGTCTTCTGG 546
Db 2 TTTGGCACCCCATGAACCCCAATCTCCAGAGAGCGGTGTGTTTCAATAGTCTTACTGG 61

Qy 547 AAACCTACAGATGGGGCTGGAATGATGTTATCTGTGAACTAGAAGGAATTCATATGT 606
Db 62 ATNCTTCGAATGGGGCTGGAATGATGTTTCTGTGATAGTAAACACATTCATATGT 121

Qy 607 GAGATGAATAAGATTTACCTA 627
Db 122 GAAATGAAGAAGATTTACCTA 142

RESULT 7
US-08-772-440-40
; Sequence 40, Application US/08772440
; Patent No. 6046158
; GENERAL INFORMATION:
; APPLICANT: Ariizumi, Kiyoshi

```



```

; APPLICANT: Takashima, Akira
; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE
; TITLE OF INVENTION: LECTINS, DECTIN-1 AND DECTIN-2; COMPOSITIONS AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/772,440
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTXD:493
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 152 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-772-440-40

Query Match 15.1%; Score 94.4; DB 3; Length 152;
Best Local Similarity 78.5%; Pred. No. 7.9e-21;
Matches 113; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Qy 226 TGGGGATGTTGCCAGCTCTTGGGAAGTCATTGTTGGTTCAGTTGCTACTTCATTTCACGT 285
Db 9 TGGGGATGTCGCCAAATCACTGGGAAGTCATTGTTGGTTCAGCTGCTACCTCATTTCTACC 68

Qy 286 GAAGAGAAGGTTGGTCTTAAGAGTCAGCAGAACTGTGTTGAGATGGGAGCACATTTGGTT 345
Db 69 AAGGAGAAGCTTCTGGAGCACCACTGAGCAGAACTGTGTTGAGATGGGGGCTCATCTGGTG 128

Qy 346 GTGTTCAACAGACAGCAGCAG 369
Db 129 GTGATCAATCACTGAAGCGGAGCAG 152

RESULT 8
US-08-772-440-33
; Sequence 33, Application US/08772440
; Patent No. 6046158
; GENERAL INFORMATION:
; APPLICANT: Ariizumi, Kiyoshi
; APPLICANT: Takashima, Akira
; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE
; TITLE OF INVENTION: LECTINS, DECTIN-1 AND DECTIN-2; COMPOSITIONS AND USES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```





; NAME/KEY: modified\_base  
; LOCATION: 3405..6871  
; OTHER INFORMATION: /mod\_base= OTHER  
; OTHER INFORMATION: /note= "S = C or G"  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: 3457..9998  
; OTHER INFORMATION: /mod\_base= OTHER  
; OTHER INFORMATION: /note= "W = A or T"  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: 3595..9999  
; OTHER INFORMATION: /mod\_base= OTHER  
; OTHER INFORMATION: /note= "Y = C or T"  
US-08-772-440-33

Query Match 8.3%; Score 52; DB 3; Length 10409;  
Best Local Similarity 80.3%; Pred. No. 2.1e-06;  
Matches 61; Conservative 0; Mismatches 15; Indels 0; Gaps 0;  
QY 294 GGTTTGGCTTACAGTGGACGACAACTGTTGAGATGGGAGCACATTTGGTTGTTTCAA 353  
Db 8777 GGCTTGGACACAGCTGAGCAGAACTGTGTCAGATGGGGGCTCATCTGGTGTGATCAA 8718  
QY 354 CACAGAAGCAGACGAG 369  
Db 8717 TACTGAAGCGGAGCAG 8702

RESULT 14  
US-08-650-578-1  
; Sequence 1, Application US/08650578  
; Patent No. 5811284  
; GENERAL INFORMATION:  
; APPLICANT: Chang, Chiwen  
; APPLICANT: Aramburu Beltran, Jose  
; APPLICANT: Lopez-Botet, Miguel  
; APPLICANT: Phillips Jr., Joseph H.  
; APPLICANT: Lanier, Lewis L.  
; TITLE OF INVENTION: Purified Mammalian NK Antigens and  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: DNAX Research Institute  
; STREET: 901 California Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94304-1104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/650,578  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/175,339  
; FILING DATE: 29-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ching, Edwin P.  
; REGISTRATION NUMBER: 34,090  
; REFERENCE/DOCKET NUMBER: DX0391  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-852-9196  
; TELEFAX: 415-496-1200  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 871 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single

; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 122..661  
US-08-650-578-1  
Query Match 7.0%; Score 44.2; DB 1; Length 871;  
Best Local Similarity 53.1%; Pred. No. 0.00019;  
Matches 94; Conservative 0; Mismatches 83; Indels 0; Gaps 0;  
QY 201 CAGTGAAGGACAAAGTGCACGCTGGGATGTTGCCACAGCTTCTTGGAAAGTCATTGCG 260  
Db 268 CATAGAACTCCAGAAAGACTCTGACTGCTGTTCTTGCCAAAAAATGGGTGGGTACCG 327  
QY 261 TTCACAGTTGCTACTTTCATTTCCAGTGAAGAGAAGGTTTGGTCTAAGAGTGAGCAGAACTG 320  
Db 328 GTGCAACTGTACTTTCATTTCCAGTGAACAGAAAACCTTGAACGAAAGTCGGCATCTCTG 387  
QY 321 TGTGTGAGATGGGACACATTTGGTTGTGTTTCAACACAGAAGCAGACAGACAGAAATTCAT 377  
Db 388 TGCTTCTCAGAAAAATCCAGCCTGCTTCAGCTTCAAAACACAGATGAATGGATTTTAT 444  
RESULT 15  
US-09-222-575-29/c  
; Sequence 29, Application US/09222575  
; Patent No. 6387697  
; GENERAL INFORMATION:  
; APPLICANT: Yugu, Jiang  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Xu, JIangchun  
; TITLE OF INVENTION: Compositions for the Treatment and Diagnosis of Breast Cancer  
; FILE REFERENCE: 210121.470  
; CURRENT APPLICATION NUMBER: US/09/222,575  
; CURRENT FILING DATE: 1998-12-28  
; NUMBER OF SEQ ID NOS: 174  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 29  
; LENGTH: 301  
; TYPE: DNA  
; ORGANISM: Human  
US-09-222-575-29  
Query Match 6.4%; Score 40.2; DB 4; Length 301;  
Best Local Similarity 53.1%; Pred. No. 0.0022;  
Matches 112; Conservative 0; Mismatches 93; Indels 6; Gaps 1;  
QY 408 TTTTTCGGGGCTTTCAGACCCCAAGGTAATAATAATTGGCAATGGATTGATAAGACACC 467  
Db 301 TTTTATTGGACTGTCAGACCAGGTTGTCGAGGGTCAATGGCAATGGGTGGACGCACACC 242  
QY 468 TTATGACAAAAATGTCAGATTTTGGCACCTAGCTAGCCCAATCTCTGCA-----GA 521  
Db 241 TTTGACAAAGTCTCTGAGCTTCTGGGATGTAGGGGAGGCCCAACAACATAGCTACCTGGA 182  
QY 522 GCAATGTGCTTCAATAGTCTTCTGGAAAACCTACAGGATGGGCTGGAATCATGTTATCTG 581  
Db 181 GGACTGTGCCACCATGAGAGACTCTTCAAAACCCCAAGGCAAAATTTGCAATGATGTAACCTG 122  
QY 582 TGAAGTGAAGGAATTCATATGTGAGATG 612  
Db 121 TTTCTCAATTATTTTCGGGATTTGTGAAATG 91

Search completed: February 17, 2003, 12:10:44  
Job time : 29.425 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 17, 2003, 08:44:21 : Search time 90.6728 Seconds  
(without alignments)  
15572.506 Million cell updates/sec

Title: US-09-766-511B-52

Perfect score: 627

Sequence: 1 atgatgcagagcagcaacc.....agatgaataagattaccta 627

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002.\*

- 1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*
- 2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*
- 3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*
- 4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*
- 5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*
- 6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*
- 7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*
- 8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*
- 9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*
- 10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*
- 11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*
- 12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*
- 13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*
- 14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*
- 15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*
- 16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*
- 17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*
- 18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*
- 19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*
- 20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*
- 21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*
- 22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*
- 23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*
- 24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	627	100.0	3114	22	AAS01375	Human TANGO 405 cd
2	484.4	77.3	850	20	AAZ07531	Human SDCMP3 polyp
3	373.2	59.5	630	20	AAZ07532	Mouse SDCMP3 polyp
4	373.2	59.5	1227	19	AAV42549	Mouse dectin-2 cDN
5	373.2	59.5	1252	22	AAS01378	Murine TANGO 405 a
6	344.8	55.0	821	22	AAS01376	Murine TANGO 405 c
7	303.2	48.4	501	19	AAV42554	Mouse dectin-2 ext
8	273.4	43.6	1312	22	AAF90241	Nucleotide sequenc
9	269.4	43.0	827	22	AAD19729	Dendritic cell (DC

10	265.6	42.4	800	22	AAD19730	Dendritic cell (DC
11	207.2	33.0	324	22	AAK63162	Human immune/haema
12	201.2	32.1	444	24	ABK52914	Human cDNA encodin
13	200.4	32.0	402	24	ABK52903	Human cDNA encodin
14	192	30.6	1091	22	AAS31385	Human cDNA encodin
15	192	30.6	1091	24	ABQ66709	Human polynucleoti
16	192	30.6	1096	22	AAS31224	Human cDNA encodin
17	192	30.6	1096	24	ABQ66548	Human polynucleoti
18	192	30.6	1104	20	AAK04855	Primate DCMP1 C-le
19	192	30.6	1307	24	ABU90686	Human polynucleoti
20	176	28.1	582	24	ABK52904	Human dendritic ce
21	172.8	27.6	753	24	ABK52906	Human dendritic ce
22	166.8	26.6	1013	24	ABK52905	Human dendritic ce
23	151.4	24.1	558	24	ABK52907	Human dendritic ce
24	141.8	22.6	561	24	ABK52911	Human dendritic ce
25	135	21.5	448	24	ABK52912	Human dendritic ce
26	135	21.5	666	24	ABK52910	Human dendritic ce
27	134	21.4	568	24	ABK52908	Human dendritic ce
28	134	21.4	1418	20	AAK04867	rodent DCMP1 C-lec
29	130.2	20.8	403	24	ABK52909	Human dendritic ce
30	117.4	18.7	968	22	AAF81743	Human membrane ass
31	117.4	18.7	997	20	AAS52274	Protein PRO244 cDN
32	117.4	18.7	997	22	AAS45936	Human DNA encoding
33	117.4	18.7	997	22	AAF72432	Human PRO244 cDNA
34	105.6	16.8	323	22	AAK53734	Murine transport a
35	96.6	15.4	559	24	ABK52913	Human dendritic ce
36	96	15.3	449	22	AAK63968	Human immune/haema
37	94.4	15.1	10409	19	AAV42558	Mouse dectin-2 gen
38	91	14.5	2517	22	AAF86349	Murine cDNA encodi
39	77	12.3	2059	21	AAA26386	Human secreted pro
40	60	9.6	2256	22	AAH43054	Nucleotide sequenc
41	59.8	9.5	1521	23	AAK71133	DNA encoding novel
42	59.8	9.5	2005	24	ABL95574	Human angioogenesis
43	59.8	9.5	2005	24	ABL88085	Human PRO7223 cDNA
44	59.8	9.5	2024	21	AAA07697	Human collectin en
45	59.8	9.5	2181	22	AAI60628	Human polynucleoti

ALIGNMENTS

RESULT 1  
AAS01375  
ID AAS01375 standard; cDNA; 3114 BP.

XX AAS01375;

XX 04-JUL-2001 (first entry)

DT Human TANGO 405 cDNA sequence.

DE

XX Human; TANGO 210; clone jthLal52h06; TANGO 364; TANGO 366; dectin-2;  
INTERCEPT 394; INTERCEPT 400; TANGO 405; cellular process regulator;  
gene therapy; growth modulator; proliferation; cell differentiation;  
lymphocyte; bone marrow cell migration; leukaemia; lymphoma;  
autoimmune disorder; ss.

XX Homo sapiens.

OS Key Location/Qualifiers

FT CDS 154..783

FT FT /\*tag= a

FT FT /product= "TANGO 405 protein"

FT FT /note= "The ORF is specifically claimed"

FT FT sig\_peptide 154..297

FT FT mat\_peptide /\*tag= b

FT FT /\*tag= c

XX WO200118016-A1.

XX 15-MAR-2001.

PF 30-JUN-2000; 2000WO-US18174.  
XX 10-SEP-1999; 99US-0393996.  
XX (MILL-) MILLENNIUM PHARM INC.  
XX Fraser CC, Sharp JD, Wrighton N, Myers PS, Goodearl ADJ;  
XX WPI: 2001-183280/18.  
XX P-PSDB; AAU00479.  
XX Isolated nucleic acid molecules encoding proteins useful as modulating  
PT agents in regulating a variety of cellular processes are used for  
PT treating e.g. cancer and autoimmune disorders -  
XX Claim 2; Fig 6A-6C; 326pp; English.  
XX The present sequence encoding for human TANGO 405 protein is isolated  
CC from cDNA clone jthlal52h06 from a human mixed lymphocyte reaction cDNA  
CC library. It is 1 of 6 novel human proteins which include TANGO 210  
CC (AAU00469), TANGO 364 (AAU00471), TANGO 366 (AAU00472), INTERCEPT 394  
CC (AAU00473), and INTERCEPT 400 (AAU00476). Novel sequences for murine  
CC TANGO 210 (AAU00470), INTERCEPT 400 (AAU00477), TANGO 405 (AAU00480) and  
CC a rat INTERCEPT 400 (AAU00478) sequence are also described. The nucleic  
CC acids encoding these novel proteins are useful as modulating agents in  
CC regulating a variety of cellular processes and can be used to express  
CC the proteins in a host cell in gene therapy applications. Human and  
CC murine TANGO 405 proteins show sequence homology to murine lectin-2.  
CC TANGO 405 modulates growth, proliferation, survival, differentiation,  
CC activity, morphology and movement/migration of human lymphocytes and  
CC bone marrow cells and tissues and can be used to prevent, diagnose or  
CC treat leukaemia, lymphomas and autoimmune disorders.  
XX SQ Sequence 3114 BP; 1001 A; 527 C; 517 G; 1069 T; 0 other;  
Query Match 100.0%; Score 627; DB 22; Length 3114;  
Best Local Similarity 100.0%; Pred. No. 6.4e-193;  
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGATGCAAGAGCAGCAACCTCAAAGTACAGAGAAAGAGCGTGTTCCTCGAGACTC 60  
DB 154 ATGATGCAAGAGCAGCAACCTCAAAGTACAGAGAAAGAGCGTGTTCCTCGAGACTC 213  
QY 61 TGGTCTGTGGCTGGGATTTCCATTGCACTCCTCAGTGTCTTCAATGTGAGCTGTGA 120  
DB 214 TGGTCTGTGGCTGGGATTTCCATTGCACTCCTCAGTGTCTTCAATGTGAGCTGTGA 273  
QY 121 GTAACCTTACCATTTTACATATGGTGAACCTGGCAAAAGGCTGTCTGAACCTACACTCATAT 180  
DB 274 GTAACCTTACCATTTTACATATGGTGAACCTGGCAAAAGGCTGTCTGAACCTACACTCATAT 333  
QY 181 CATTCAAGTCTCACCTGCTTCAGTGAAGGACAAAGGTGCCAGCCTGGGGATGTTGCCCA 240  
DB 334 CATTCAAGTCTCACCTGCTTCAGTGAAGGACAAAGGTGCCAGCCTGGGGATGTTGCCCA 393  
QY 241 GCTTCTTGGAGTCATTTGGTTCCAGTTGCTACTTCAATTTCCAGTGAAGAGAAGTTGG 300  
DB 394 GCTTCTTGGAGTCATTTGGTTCCAGTTGCTACTTCAATTTCCAGTGAAGAGAAGTTGG 453  
QY 301 TCTAAGAGTGAAGCAACTGCTTGAGATGGGACCACATTTGCTGTGTTCAACACAGAA 360  
DB 454 TCTAAGAGTGAAGCAACTGCTTGAGATGGGACCACATTTGCTGTGTTCAACACAGAA 513  
QY 361 GCAGAGCAGAAATTTCAATGTCCAGCAGCTGAATGAGTCATTTTCTTATTTCTTGGGGCTT 420  
DB 514 GCAGAGCAGAAATTTCAATGTCCAGCAGCTGAATGAGTCATTTTCTTATTTCTTGGGGCTT 573  
QY 421 TCAGACCCACAGGTAATATAATTTGCCAATGGATTGATTAAGACACCTTATGAGAAAAAT 480  
DB 574 TCAGACCCACAGGTAATATAATTTGCCAATGGATTGATTAAGACACCTTATGAGAAAAAT 633  
QY 481 GTCAGATTTTGGCACCTAGGTGAGCCCAATCATTTCTGCAGAGCAATGTGCTTCAATAGTC 540

DB 634 GTCAGATTTTGGCACCTAGGTGAGCCCAATCATTTCTGCAGAGCAATGTGCTTCAATAGTC 693  
QY 541 TTCTGGAACCTACAGAGTGGGCTGGGAATGATGTTATCTCTGAACTAGAGAATTCA 600  
DB 694 TTCTGGAACCTACAGAGTGGGCTGGGAATGATGTTATCTCTGAACTAGAGAATTCA 753  
QY 601 ATATGTGAGATGAATAAGATTTTACCTA 627  
DB 754 ATATGTGAGATGAATAAGATTTTACCTA 780  
RESULT 2  
AAZ07531  
ID AAZ07531 standard; cDNA; 850 BP.  
XX AC  
XX AAZ07531;  
XX DT 26-NOV-1999 (first entry)  
DE Human SDCMP3 polypeptide encoding cDNA.  
XX KW Binding compound; antibody binding site; primate; rodent; SDCMP3; SDCMP4;  
KW Schering dendritic cell membrane protein; dendritic cell physiology;  
KW genetic fingerprinting; cancer immunotherapy; abnormal proliferation;  
KW cancer; forensic; human; lectin 73; ss.  
XX OS Homo sapiens.  
FH Key Location/Qualifiers  
FT CDS 108..596  
FT /\*tag= a  
FT /product= "SDCMP3"  
XX PN WO9947673-A2.  
PD 23-SEP-1999.  
XX PF 16-MAR-1999; 99WO-US03740.  
XX PR 17-MAR-1998; 98US-0040111.  
XX (SCHE ) SCHERING CORP.  
PA Chalus L, Quan AB, Bates EEM, Gorman DM, Saeland S, Lebecque SJE;  
PI Philipps JH;  
XX WPI: 1999-562114/47.  
DR P-PSDB; AAY27446.  
XX Binding compound specific for primate or rodent Schering dendritic cell  
PT membrane proteins -  
XX Claim 8; Page 80; 89pp; English.  
XX The invention relates to a binding compound comprising an antibody  
CC binding site which specifically binds to: (a) a primate or rodent  
CC Schering dendritic cell membrane protein 3 (SDCMP3) protein; or (b) a  
CC primate SDCMP4 protein. The binding compound can be used to modulate  
CC dendritic cell physiology or function. The sequences may also be used as  
CC probes in forensic techniques, such as genetic fingerprinting. They can  
CC also be used to distinguish tissue and cell types in situ or in vitro.  
CC The sequences may also be useful in cancer immunotherapy and in the  
CC treatment of conditions associated with abnormal physiology or  
CC or degenerative conditions. The present sequence represents a cDNA  
CC encoding the human SDCMP3 (lectin 73).  
SQ Sequence 850 BP; 257 A; 153 C; 188 G; 252 T; 0 other;  
Query Match 77.3%; Score 484.4; DB 20; Length 850;  
Best Local Similarity 98.8%; Pred. No. 7.5e-147;  
Matches 488; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy	1	ATGATGCAAGAGCAGCAACCTCAAAGTACAGAGAAAAGAGGCTGGTTGTCCTCCTGAGACTC	60
Db	108	ATGATGCAAGAGCAGCAACCTCAAAGTACAGAGAAAAGAGGCTGGTTGTCCTCCTGAGACTC	167
Qy	61	TGGTCTCTGGCTGGGATTTCCATTGCGACTCCTCAGTGTGCTTCATTGTGAGCTGTGTA	120
Db	168	TGGTCTCTGGCTGGGATTTCCATTGCGACTCCTCAGTGTGCTTCATTGTGAGCTGTGTA	227
Qy	121	GTAACCTTACCATTTTACATATGGTGAACCTGGCAAAAGGCTGCTGAACCTACACTCATAT	180
Db	228	GTAACCTTACCATTTTACATATGGTGAACCTGGCAAAAGGCTGCTGAACCTACACTCATAT	287
Qy	181	CATTCAAGTCTCACTGCTTCCAGTGAAGGACAAAGGTGCCAGCTGGGGATGTGGCCCA	240
Db	288	CATTCAAGTCTTACCTGCTTCCAGTGAAGGACAAAGGTGCCAGCTGGGGATGTGGCCCA	347
Qy	241	GCTTCTTGGAGTCAATTTGGTTCAGTTGCTACTTTCATTTCCAGTGAAGAGGATTTGG	300
Db	348	GCTTCTTGGAGTCAATTTGGTTCAGTTGCTACTTTCATTTCCAGTGAAGAGGATTTGG	407
Qy	301	TCTAAGAGTGAGCAGAACTGTGTTGAGATGGGACACATTTGGTTGTTCACACAGAA	360
Db	408	TCTAAGAGTGAGCAGAACTGTGTTGAGATGGGACACATTTGGTTGTTCACACAGAA	467
Qy	361	GCAGAGCAGAAATTCATTGTCAGCAGCTGAATCAGTCAATTTCTTATTTCTGGGGCTT	420
Db	468	GCAGAGCAGAAATTCATTGTCAGCAGCTGAATCAGTCAATTTCTTATTTCTGGGGCTT	527
Qy	421	TCAGACCCACAGGTAATAATAATTTGGCAATGGATTGATTAAGACACCTTATGAGAAAAAT	480
Db	528	TCAGACCCACAGGTAATAATAATTTGGCAATGGATTGATTAAGACACCTTATGAGAAAAAT	587
Qy	481	GTCAGATTTTGGCA	494
Db	588	GTCAGGTGAGTGCA	601
RESULT 3			
ID	AAZ07532 standard; cDNA; 630 BP.		
XX	AAZ07532;		
AC	AAZ07532;		
DT	26-NOV-1999 (first entry)		
XX	Mouse SDCMP3 polypeptide encoding cDNA.		
DE			
XX			
KW	Binding compound; antibody binding site; primate; rodent; SDCMP3; SDCMP4;		
KW	Schering dendritic cell membrane protein; dendritic cell physiology;		
KW	genetic fingerprinting; cancer immunotherapy; abnormal proliferation;		
KW	cancer; forensic; mouse; ss.		
XX			
OS	Mus sp.		
XX			
PH	Key	Location/Qualifiers	
FT	CDS	1..630	
FT		/*tag= a	
FT		/product= "SDCMP3"	
XX			
PN	W09947673-A2.		
XX			
PD	23-SEP-1999.		
XX			
XX	16-MAR-1999; 99WO-US03740.		
PR			
PR	17-MAR-1998; 98US-0040111.		
XX			
PA	(SCHE ) SCHERING CORP.		
XX			
PI	Chalus L, Quan AB, Bates EEM, Gorman DM, Saeland S, Lebecque SJE;		
PI	Phillips JH;		
XX			
DR	WPI; 1999-562114/47		

DR	P-PSDB; AAY27447.		
XX			
PT	Binding compound specific for primate or rodent Schering dendritic cell		
PT	membrane proteins		
XX			
PS	Claim 8; Page 82; 89pp; English.		
XX			
CC	The invention relates to a binding compound comprising an antibody		
CC	binding site which specifically binds to: (a) a primate or rodent		
CC	Schering dendritic cell membrane protein 3 (SDCMP3) protein; or (b) a		
CC	primate SDCMP4 protein. The binding compound can be used to modulate		
CC	dendritic cell physiology or function. The sequences may also be used as		
CC	probes in forensic techniques, such as genetic fingerprinting. They can		
CC	also be used to distinguish tissue and cell types in situ or in vitro.		
CC	The sequences may also be useful in cancer immunotherapy and in the		
CC	treatment of conditions associated with abnormal physiology or		
CC	development, including abnormal proliferation, e.g. cancerous conditions,		
CC	or degenerative conditions. The present sequence represents a cDNA		
XX	encoding the mouse SDCMP3.		
SQ	Sequence 630 BP; 171 A; 135 C; 152 G; 172 T; 0 other;		
Query Match 59.5%; Score 373.2; DB 20; Length 630;			
Best Local Similarity 76.1%; Pred. No. 9.2e-111;			
Matches 475; Conservative 0; Mismatches 143; Indels 6; Gaps 1;			
QY	10	GAGCACAACCTCAAGTACAGAGAAAAGAGGCTGGTTGTCCTCCTCAGACTCTGGTCTGTG	69
Db	4	GTGCGAAGAACAAATCCCAAGGAAGGAGTCTGCTGGACCTCAGACTCTGGTCACTG	63
QY	70	GCTGGATTTCATTCACCTCCTCAGTCTGCTTCATTGTGAGCTGTCTAGTAACTTAC	129
Db	64	GCTGTGATTTCATTCATCTTGTAGTACTGTTTCATTGCGAGCTGTGGTGACTTAC	123
QY	130	CATTTTACATATGGTGAACACTGGCAAAAGGCTGCTGAACACTACATCATATCAATCAAGT	189
Db	124	CAATTTATTATGACCCAGCCAGTATGAACTATATGAACACTTACACATACATCCAGT	183
QY	190	CTCACCTGCTCAGTGAAGGGACAAAGTGGCCAG-----CCTGGGGATGTTGCCCAGCT	243
Db	184	CTCACCTGCTCAGTGAAGGGACTATGGTGTGAGAAAATGTGGGGATGCTGCCCAAT	243
QY	244	TCTTGAAGTCAATTTGGTTCAGTGTCTACTTCATTTCCTCCAGTGAAGAGAGTGTGGTCT	303
Db	244	CACCTGAAGTCAATTTGGTTCAGTGTCTACTTCATTTCCTCCAGTGAAGAGAGTGTGGGAC	303
QY	304	AGAGTGACGACAACTGTGTTGAGATGGAGCACATTTGGTTGTGTTCAACACAGAAAGCA	363
Db	304	ACCAGTGACGACAACTGTGTTGAGATGGAGCACATTTGGTTGTGTTCAACACAGAAAGCA	363
QY	364	GAGCAGAAATTCATTCGCCAGCAGCTGAATGAGTCAATTTCTTATTTCTGGGGCTTTCA	423
Db	364	GAGCAGAAATTCATTCGCCAGCAGCTGAATGAGTCAATTTCTTATTTCTGGGGCTTTTCG	423
QY	424	GACCCACAAGTAAATAATTTGGCAATGGATTGATTAAGACACACCTTATGAGAAAAATGTC	483
Db	424	GATCCACAAGTAAATGGCAATGGCAATGGATCGATGATACCTTCCTTCAGTCAAAATGTC	483
QY	484	AGATTTTGGCACCCTAGGTGAGCCCAATCATTTCTGAGAGCAATGTGTTCAATAGTCTTC	543
Db	484	AGGTTCCTGCGACCCCCATGACCCCAATCTTCCAGAGAGCGGTGTGTTTCAATAGTTTAC	543
QY	544	TGGAACCTFACAGGATGGGCTGGAAATGATGTTATCTGTGAACCTAGAAGGAATCAATA	603
Db	544	TGGAATCTTCGAAATGGGGCTGGAAATGATGTTTCTGTGATAGTAAACACAAATCAATA	603
QY	604	TGTGAGATGAATAAGATTTACCTA	627
Db	604	TGTGAAATCAAGAAGATTTACCTA	627
RESULT 4			
AAV42549			





PI Fraser CC, Sharp JD, Wrighton N, Myers PS, Goodearl ADJ;  
DR WPI: 2001-183280/18.  
DR P-PSDB; AAU00482.  
XX  
PT Isolated nucleic acid molecules encoding proteins useful as modulating  
PT agents in regulating a variety of cellular processes are used for  
PT treating e.g. cancer and autoimmune disorders -  
XX  
PS Disclosure; Fig 6N-6P; 326pp; English.  
XX  
CC The present sequence encoding for murine TANGO 405 alternative splice  
CC variant protein is isolated from cDNA clone jtmMa025all from a long-term  
CC bone marrow cDNA library. TANGO 405 (AAU00480) is 1 of 3 novel murine  
CC proteins which include TANGO 210 (AAU00470) and INTERCEPT 400  
CC (AAU00477). Six novel human proteins which include TANGO 210 (AAU00469),  
CC TANGO 364 (AAU00471), TANGO 366 (AAU00472), INTERCEPT 394 (AAU00473),  
CC INTERCEPT 400 (AAU00476) and TANGO 405 (AAU00479), and a rat  
CC INTERCEPT 400 (AAU00478) sequence are also described. The nucleic acids  
CC encoding these novel proteins are useful as modulating agents in  
CC regulating a variety of cellular processes and can be used to express  
CC the proteins in a host cell in gene therapy applications. Human and  
CC murine TANGO 405 proteins show sequence homology to murine dectin-2.  
CC TANGO 405 modulates growth, proliferation, survival, differentiation,  
CC activity, morphology and movement/migration of human lymphocytes and  
CC bone marrow cells and tissues and can be used to prevent, diagnose or  
CC treat leukaemia, lymphomas and autoimmune disorders.  
XX  
SQ Sequence 1252 BP; 330 A; 283 C; 286 G; 353 T; 0 other;

Query Match 59.5%; Score 373.2; DB 22; Length 1252;  
Best Local Similarity 76.1%; Pred. No. 1.3e-110;  
Matches 475; Conservative 0; Mismatches 143; Indels 6; Gaps 1;

Qy 10 GAGCAGCAACCTCAAGTACAGAGAAAGGCTGTGTCCTGAGACTCTGGTCTGTG 69  
Db 182 GTGCAGGAAAGACATCCCAAGGAGGAGTCTGTGGACCTGAGACTCTGGTCA 241  
Qy 70 GCTGGGATTTCCATTGCTCACCTCCTCAGTGTCTTCAATTTGAGTCTGTAGTAACTTAC 129  
Db 242 GCTGTGATTTCCATTGCTCACCTCCTCAGTGTCTTCAATTTGAGTCTGTAGTAACTTAC 301  
Qy 130 CATTTTACATATGCTGAACTGGCAAAAGGCTGTCTGAACTACACTCATATCATTCAGT 189  
Db 302 CAATTTATTTGACACAGCCAGTGAAGACTATATGAACCTTACACATACCATTCAGT 361  
Qy 190 CTCACCTGCTTCAAGGAGCAAAAGTCCAG-----CCTGGGGAGTGTGCCAGCT 243  
Db 362 CTCACCTGCTTCAAGGAGCAATGCTGTGCAAAAAATGTGGGGATGCTGCCCAAT 421  
Qy 244 TCTTGAAGTCATTTGCTTCCAGTGTCTACTTCAATTTCCAGTCAAGAGGTTTGGTCT 303  
Db 422 CACTGGAAGTCATTTGGCTCCAGTGTCTACTTCTACCTATTCTACCAAGGAGAACTTCTGGAGC 481  
Qy 304 AAGAGTGAGCAGAACTGTGTTGAGATGGGAGCACATTTGTTTCAACAGAGCA 363  
Db 482 ACCAGTGAGCAGAACTGTGTTGAGATGGGGCTCATCTGGTGTGATCAATATCAAGCG 541  
Qy 364 GAGCAGAAATTTCAATGTCACAGCTGAATGAGTCAATTTTATTTTCGCGGCTTTCA 423  
Db 542 GAGCAGAAATTTCAATCACCAGCAGCTGAATGAGTCAATTTTCTTACTTCTCGGTCTTCG 601  
Qy 424 GACCCACAAGGTATATATATGTCATGATGATTCATTAAGACACCTTATGAGAAATGTC 483  
Db 602 GATCCACAAGGTATGCAATGCAATGGATGATGATGATGATGATGATGATGATGATGATG 661  
Qy 484 AGATTTTGGCACCTTAGTGGCCCAATCATCTCTGACAGCAATGTGCTTCAATAGTCTTC 543  
Db 662 AGGTTCTGGCACCCCATGAACCAATCTTCCAGAGAGCGGTGTGTTTCAATAGTCTTAC 721  
Qy 544 TGGAAACCTACAGATGGGCTGGAATGATGTTATCTGTGAAACTAGAAGGAATTCATA 603  
Db 722 TGGAACTCTTGGAAATGGGCTGGAATGATGTTTCTTCTGTCATAGTAAACACAAATTCATA 781

Qy 604 TCTGACATGAATAAGATTACCTA 627  
Db 782 TGTGAATGAAGAAGATTACCTA 805

RESULT 6  
AAS01376  
ID AAS01376 standard; cDNA; 821 BP.  
XX  
AC AAS01376;  
XX  
DT 04-JUL-2001 (first entry)  
XX  
DE Murine TANGO 405 cDNA sequence.  
XX  
KW Murine; TANGO 210; clone jtmMa025all; TANGO 364; TANGO 366; dectin-2;  
KW INTERCEPT 394; INTERCEPT 400; TANGO 405; cellular process regulator;  
KW gene therapy; growth modulator; proliferation; cell differentiation;  
KW lymphocyte; bone marrow cell migration; leukaemia; lymphoma;  
KW autoimmune disorder; ss.  
XX  
OS Mus sp.  
XX  
FH Key Location/Qualifiers  
FT 174..710  
FT CDS /\*tag= a  
FT /product= "TANGO 405 protein"  
FT sig\_peptide 174..299  
FT /\*tag= b  
FT mat\_peptide 300..707  
FT /\*tag= c  
XX  
PN WO200118016-A1.  
XX  
PD 15-MAR-2001.  
XX  
PF 30-JUN-2000; 2000WO-US18174.  
XX  
PR 10-SEP-1999; 99US-0393996.  
XX  
PA (MILL-) MILLENNIUM PHARM INC.  
XX  
PI Fraser CC, Sharp JD, Wrighton N, Myers PS, Goodearl ADJ;  
XX  
DR WPI: 2001-183280/18.  
DR P-PSDB; AAU00480.  
XX  
PT Isolated nucleic acid molecules encoding proteins useful as modulating  
PT agents in regulating a variety of cellular processes are used for  
PT treating e.g. cancer and autoimmune disorders -  
XX  
PS Disclosure; Fig 6E-6F; 326pp; English.  
XX  
CC The present sequence encoding for murine TANGO 405 protein is isolated  
CC from cDNA clone jtmMa025all from a long-term bone marrow cDNA library.  
CC An alternative splice variant of this protein is also described  
CC (AAU00482). Other novel murine proteins include TANGO 210 (AAU00470) and  
CC INTERCEPT 400 (AAU00477). Six novel human proteins which include  
CC TANGO 210 (AAU00469), TANGO 364 (AAU00471), TANGO 366 (AAU00472),  
CC INTERCEPT 394 (AAU00473), INTERCEPT 400 (AAU00476) and TANGO 405  
CC (AAU00479), and a rat INTERCEPT 400 (AAU00478) sequence are also  
CC described. The nucleic acids encoding these novel proteins are useful as  
CC modulating agents in regulating a variety of cellular processes and can  
CC be used to express the proteins in a host cell in gene therapy  
CC applications. Human and murine TANGO 405 proteins show sequence homology  
CC to murine dectin-2. TANGO 405 modulates growth, proliferation, survival,  
CC differentiation, activity, morphology and movement/migration of human  
CC lymphocytes and bone marrow cells and tissues and can be used to prevent,  
CC diagnose or treat leukaemia, lymphomas and autoimmune disorders.  
XX  
SQ Sequence 821 BP; 214 A; 188 C; 195 G; 223 T; 1 other;

PA	(TEXA ) UNIV TEXAS SYSTEM.								
XX	Ariizumi K, Takashima A;								
XX	WPI: 1998-377594/32.								
DR	P-PSDB; AAW63022.								
XX	Nucleic acid encoding dendritic cell specific peptide(s) dectin-1								
PT	and -2 - useful, e.g. to regulate immune response, as vaccine								
PT	adjuvants, for diagnosis and drug screening								
XX	Disclosure: Page 153; 200pp: English.								
PS	This cDNA sequence codes for the extracellular domain (see								
CC	AAW63022) of a novel murine dendritic cell (DC) specific member of								
CC	the C-type lectin family, termed DC-associated C-type lectin-2, or								
CC	dectin-2 (see also AAW63010), that is essential for DC-mediated T								
CC	cell activation. Full-length cDNA (see AAV42549) was isolated by								
CC	subtractive hybridisation from BALB/c mouse epidermis DC line XS52.								
CC	The invention provides: dectin-1 and -2 polypeptides (see also								
CC	AAW63009-22 and AAW69236-37), useful for purifying T cells, for								
CC	detecting autoantibodies and for up-regulating immunity e.g. as								
CC	vaccine adjuvants; dectin DNA (see AAV42548-54, AAV42558-60 and								
CC	AAV44850-51); expression vectors; recombinant host cells; probes								
CC	and primers; antibodies; compounds that modulate dectin-mediated								
CC	activation of T cells; transgenic animals; and dectin ligands.								
CC	Dectin expression can also be down-regulated to treat allergy and								
CC	autoimmune disease. Recombinant extracellular domain has been								
CC	expressed as a His-tagged protein (see AAW69237) useful for antibody								
CC	production.								
XX	Sequence 501 BP; 142 A; 108 C; 115 G; 136 T; 0 other;								
SQ									
Query Match	48.4%;	Score 303.2;	DB 19;	Length 501;					
Best Local Similarity	78.4%;	Pred. No. 4.7e-88;							
Matches	378;	Conservative 0;	Mismatches 98;	Indels 6;	Gaps 1;				
QY	152	GCAAAAGGCTGTGTGAACCTACACTCATATCATTTCAAGTCTCACCTGCTTCAGTGAAGGGA	211						
Db	20	GTAGAAGACTATATGAACCTTCACATACCATCCAGTCTCACCTGCTTCAGTGAAGGGA	79						
QY	212	CARAAGTGGCAG-----CCTGGGGATGTGCCAGCTTCTTGGAGTCATTTGGTTCCA	265						
Db	80	CTATGGTGTGCAGAAAAAATGTGGGATGCTGCCCAAAATCACTGGAAGTCAATTTGGCTCCA	139						
QY	266	GTTGCTACTTTCATTTCCAGTGAAGAGAAGGTTTGGTCTTAAGAGTCAGCAGAACTGTGTTG	325						
Db	140	GCTGCTACCTCATTTCTACCAAGGAGAACTTCTGGAGCACCACTGAGCAGAACTGTGTTT	199						
QY	326	AGATGGGAGCACATTTGGTTGGTTTCAACACAGACAGCAGAGAAATTTCAATTTGCCAGC	385						
Db	200	AGATGGGGCTCATCTGTGGTGTATCAATACTGAAGCGGAGCAGAAATTTTCATCACCCAGC	259						
QY	386	AGCTGAATCAGTCATTTCTTATTTCTGGGCTTTCAGACCCACAAAGCTAATAATAATT	445						
Db	260	AGCTGAATGAGTCACCTTCTTACTCTCTGGGCTTTTCGGATCCACAAAGTAATGGCAAT	319						
QY	446	GGCAATGGATTTGATAAGACACACCTTATGAGAAAAATGTCAGATTTTGGCACCTAGGTGAGC	505						
Db	320	GGCAATGGATCGATGATACTCTTTTCAGTCMAAATGTCAGGTTCTGGCACCCCATGAAC	379						
QY	506	CCAATCATTTCTGCAGACAAATGTGCTTCAATAGTCTTCTTGGAAACCTCAGAGATGGGGCT	565						
Db	380	CCAATCTTCAGAAAGCGGTGTGTTTCAATAGTTTACTTGAATCCTTCCAATGGGGCT	439						
QY	566	GGAAATGATGTTATCTGTGAAACTAGAAGGAATTCATATGTGAGATGAATAAGATTACC	625						
Db	440	GGAAATGATGTTTCTGTGATAGTAAACACAATTCATATGTGAATGAAGAAGATTACC	499						
QY	626	TA 627							
Db	500	TA 501							

RESULT 8  
AAF90241  
ID AAF90241 standard; cDNA; 1312 BP.  
XX  
AC AAF90241;  
XX  
DT 06-AUG-2001 (first entry)  
XX  
DE Nucleotide sequence of BDCA-2 antigen.  
XX  
KW BDCA-2; hematopoietic cell; dendritic cell; BDCA-3; BDCA-4;  
KW viral infection; autoimmune disease; allergic response; cancer; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 491..1132  
FT FT /\*tag= a  
FT FT /product= "BDCA-2"  
XX  
XX WO200136487-A2.  
XX  
PD 25-MAY-2001.  
XX  
PF 15-NOV-2000; 2000WO-1B01832.  
XX  
PR 15-NOV-1999; 99US-0165555.  
PR 23-NOV-1999; 99US-0167076.  
PR 28-JAN-2000; 2000US-0179003.  
PR 07-FEB-2000; 2000US-0180775.  
PR 11-APR-2000; 2000US-0196824.  
PR 13-APR-2000; 2000US-0197205.  
XX  
XX (MILT-) MILTENYI BIOTECH GMBH.  
XX  
XX Schmitz J, Dzionek A, Buck DW;  
XX WPI: 2001-355622/37.  
DR P-PSDB; AAB84215.  
XX  
XX Compositions and cell populations enriched in dendritic cells through  
PT use of antigen-binding fragments specific for BDCA-2, BDCA-3 or  
PT BDCA-4, are used to treat viral infections, autoimmune diseases,  
PT allergic responses and cancer -  
XX  
PS Claim 134; Fig 12; 115pp; English.  
XX  
CC The present sequence encodes an antigen designated BDCA-2. The  
CC specification describes compositions and hematopoietic cell populations  
CC enriched in dendritic cells (DCs). These compositions are produced using  
CC antigen-binding fragments specific for BDCA-2, BDCA-3 or BDCA-4. The DCs  
CC obtained are used to treat viral infections, autoimmune disease, allergic  
CC response, and cancer. BDCA-1, BDCA-2, BDCA-3 and BDCA-4 monoclonal  
CC antibodies and their antigen-binding fragments are used to detect,  
CC enumerate and isolate DC populations from leukapheresis material, whole  
CC blood and tonsils and from non-hematopoietic and hematopoietic tissues.  
XX  
SQ Sequence 1312 BP; 374 A; 301 C; 297 G; 340 T; 0 other;

Query Match 43.6%; Score 273.4; DB 22; Length 1312;  
Best Local Similarity 66.7%; Pred. No. 3.8e-78;  
Matches 426; Conservative 0; Mismatches 201; Indels 12; Gaps 2;  
Qy 1 ATGATGCAAGAGCAGCAACCTCAAGTACAGAGAAAGAGCGTGTGTCCTCC---TGAGA 57  
Dy 491 ATGGTGCTGAAGAAGAGCCTCAAGACCGAGAGAAAGAGCTCTGGTGTCCAGTTGAAG 550  
Qy 58 CTCTGGTCTGGCTGGGATTTCCATTGCACCTCTCAGTCTGCTTCATTGTCAGCTGT 117  
Dy 551 GTCGGTCCATGGGCGTCGATCCATCTTGTCTCAGTCTGCTGTTTCTACTGTGAGTTCT 610  
Qy 118 GTAGTAACTTACCATTTTACATATGTTGTAACCTGGCAAAAGGCTGTCTGAACCTA----- 171

Db 611 GTGGTGCCTCACAATTTTATGTATAGCAAACTGTCAAGAGGCTGTCTCCAAGTTACGAGAG 670  
Qy 172 ---CACTCATATCATTTCAAGTCTCACCCTGCTTCAAGTGAAGGACAAAGGTGCCACGCTGG 228  
Db 671 TATCAACAGTATCATCCAAGGCTGACCTGGCTGCGGTCAAGTGAAGGACAAAGATAGAAGATTGG 730  
Qy 229 GGATGTTGCCAGCTTCTTTGGAAGTCATTTGGTTCCAGCTTGCTACTTTCATTTCCAGTGAA 288  
Db 731 AGCTGCTGCCCAACCCCTTGGACTTCATTTCACTAGTCTAGTCTACTTATTTCTACTGGG 790  
Qy 289 GAGAAGTTTGGTCTTAAGATGAGCAGAACTGTGTTGAGATGGGAGCAGCATTTGGTTGTG 348  
Db 791 ATGCAATCTTGACTTAAGAGTCAAAAGAACTGTTCTGTGATGGGGGCTGATCTGGTGGTG 850  
Qy 349 TTCAACACAGAAGCAGACAGAAATTTCAATTTGTCAGCAGCTCAATGAGTCATTTTCTTAT 408  
Db 851 ATCAACACAGGGAAGACAGGATTTTCATCATTCAGAACTCTGAAAAGAAATTTCTTTTAT 910  
Qy 409 TTCTCTGGGCTTTCAGAGCCCAAGGTAATAATAATTGGCAATGGATTTGATAAGACACT 468  
Db 911 TTCTCTGGGCTGTCAAGATCCAGGGGCTCGGCACATTTGGCAATGGTTTGACCAGACACCA 970  
Qy 469 TATGAGAAAAATGTCAGATTTTGGCACCTAGTGTAGTGAGGCCCAATCATTTCTGCAGAGCAATGT 528  
Db 971 TACAATGAAATGTCACATTTCTGGCACTCAGTGAACCAATAACCTTTGATGAGCGGTTGT 1030  
Qy 529 GCTTCAATAGTCTTCTGGAAACCTACAGATGGGCTGGAATGATGTTTATCTGTGAAACT 588  
Db 1031 GCGATAATAAATTTCCGTTCTTTCAAGAAATGGGCTGGAATGACATTTCACTGTGCATGTA 1090  
Qy 589 AGAAGGAATTCATATGTGAGATGAATAAGATTTTACCTA 627  
Db 1091 CCTCAGAGTCAATTTGCAAGATGAAGAGATCTACATA 1129

RESULT 9  
AAD19729  
ID AAD19729 standard; DNA; 827 BP.  
XX  
AC AAD19729;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Dendritic cell (DC) DCLEC gene.  
XX  
KW Dendritic cell; DC; DCLEC protein; gene therapy; dermatological; vaccine;  
KW atopic dermatitis; autoimmune disease; inflammatory skin disease; cancer;  
KW immunosuppressive; AIDS; Acquired immune deficiency syndrome; cytostatic;  
KW chromosomal identification; pharmaceutical; hypersensitivity; virucide;  
KW transplant rejection; chronic inflammatory disease; anti-HIV; ds.  
XX  
OS Unidentified.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..657  
FT FT /\*tag= a  
FT FT /product= "Dendritic cell DCLEC protein"  
XX  
PN WO200127773-A2.  
XX  
PD 04-OCT-2001.  
XX  
PF 28-MAR-2001; 2001WO-EP03542.  
XX  
PR 29-MAR-2000; 2000US-192934P.  
PR 18-MAY-2000; 2000US-205020P.  
PR 18-MAY-2000; 2000US-205026P.  
PR 19-MAY-2000; 2000US-205767P.  
PR 19-MAY-2000; 2000US-205769P.  
XX  
PA (NOVS ) NOVARTIS AG.  
PA (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.

XX PI Werner G, Phares W, Jaritz M, Lapp H, Kalthoff FS;  
XX WPI: 2001-616466/71.  
DR P-PSDB; AAE12079.  
XX New polypeptides for screening therapeutic agonists and antagonists  
PT comprise dendritic cell polypeptides -  
XX Claim 1; Page 50-51; 52pp; English.  
XX The invention relates to dendritic cell (DC) proteins and their  
CC corresponding DNA molecules. A pharmaceutical composition comprising  
CC agonist and antagonist of DC proteins are useful for treating abnormal  
CC conditions related to both an excess of and insufficient level of  
CC expression of DC gene, or related to both an excess of and insufficient  
CC activity of DC protein. Soluble form of DC proteins are used as an active  
CC ingredient in combination with pharmaceutical acceptable carriers.  
CC DC genes and proteins are useful for treating chronic inflammatory  
CC diseases, autoimmune diseases, transplant rejection crisis, including  
CC inflammatory skin diseases such as contact hypersensitivity, atopic  
CC dermatitis or virally-induced immune suppression such as AIDS and cancer.  
CC DC protein is useful for inducing immunological response in a mammal, and  
CC as immunogen to produce antibodies immunospecific for the polypeptide.  
CC DC gene is useful in gene therapy. DC gene is also useful as a diagnostic  
CC reagent, and for chromosomal identification. The present sequence is  
CC dendritic cell (DC) DCLEC gene which is found to belong to the family  
CC of C-type lectins with one single carbohydrate recognition domain at the  
CC C-terminal end.  
XX  
SQ Sequence 827 BP; 242 A; 164 C; 188 G; 233 T; 0 other;

Query Match 43.0%; Score 269.4; DB 22; Length 827;  
Best Local Similarity 66.8%; Pred. No. 6e-77;  
Matches 419; Conservative 0; Mismatches 196; Indels 12; Gaps 2;

QY 13 CAGCAACCTCAAGTACAGAGAAAGAGGCTGGTGTGCC---TGAGACTCTGTCTGTG 69  
DB 28 CAAGAGCCTCAAGACCGAGAGAGGAGCTCGTGGTTCACAGGTTCAGGCTCCATG 87  
QY 70 GCTGGGATTTCCATTGCTCCTCAGTCTGTCTCAATGTGAGTGTAGTAACCTTAC 129  
DB 88 CGAGTCGTATCCATCTTCTCAGTCTGTCTGTCTCAATGTGAGTGTAGTGTAGT 147  
QY 130 CATTTTACATATGGTGAACCTGGCAAAAGGCTGTCTGAACATA-----CACTCATAT 180  
DB 148 AATTTTATGTATAGCAAAACTGTCAAGAGGCTGTCCAAAGTTACGAGAGTATCAACAGTAT 207  
QY 181 CATTCAGCTCACCTGCTTCAGTGAAGGACAAAGGTGCCAGCCTGGGGATGTGCCCA 240  
DB 208 CATCAAGCCTGACCTGGCTCATGGAAGGAAGGACATAGAGATTTGGAGCTGTGCCCA 267  
QY 241 GCTTCTTGGAAGTCAATTTGGTTCAGTGTCTACTTTCATTTCAGTGAAGAGAGTTGG 300  
DB 268 ACCCTTGGACTTCATTTTCAGTCTAGTGTCTACTTTATTTCTACTGGGATCAATCTGG 327  
QY 301 TCTAAGATGAGCACTGTGTTGAGATGGGAGCACATTTGGTGTGTTCAACACAGAA 360  
DB 328 ACTAAGAGTCAAAAGCACTGTCTGTATGGGGGTGATCTGGTGGTATCAACACACAGG 387  
QY 361 CGAGGACGAATTTTCATTTGCCAGCTGAATGAGTCATTTTCTTATTTTCTGGGCTT 420  
DB 388 GAAGAACAGGATTTTCATTCATTCAGTAATCTGAAAGAAATTTCTTATTTTCTGGGCTG 447  
QY 421 TCAGACCCACAGGTAATAAATTTGGCAATGGATTGTAAGACACACCTTATGAAAAAT 480  
DB 448 TCAGATCCAGGGGTGCGGCACATTTGGCAATGGTTGACACAGACACCAATGAAAAAT 507  
QY 481 GTCATTTTGGCAGCTAGGTGAGCCCAATCATTTCTGACAGCAATGTGCTTCAATAGTC 540  
DB 508 GTCATTTTGGCAGCTAGGTGAGCCCAATCAACCTTGATGAGCGTTGTGCGCAATAAAT 567  
QY 541 TTCTGGAACCTACAGGATGGGGCTGGAAATGATGTTATCTCTGAAACAGTAAGAAATCA 600

DB 568 TTCGGTTCTTCAGAGAAATGGGGCTGGAATGACATTCACCTGTCATGTCACCTCAGAAGTCA 627  
QY 601 ATATGTGAGATGAATAAGATTTACCTA 627  
DB 628 AITTCGAAGATGAAGAAGATCTACATA 654  
RESULT 10  
AAD19730  
ID AAD19730 standard; DNA; 800 BP.  
XX  
AC AAD19730;  
DT 18-DEC-2001 (first entry)  
XX  
DE Dendritic cell (DC) DCLEC/SPLICE 1 variant gene.  
XX  
KW Dendritic cell; DC; DCLEC protein; gene therapy; dermatological; vaccine;  
KW atopic dermatitis; autoimmune disease; inflammatory skin disease; cancer;  
KW immunosuppressive; AIDS; Acquired immune deficiency syndrome; cytostatic;  
KW chromosomal identification; pharmaceutical; hypersensitivity; virucide;  
KW transplant rejection; chronic inflammatory disease; anti-HIV; variant;  
KW ds.  
XX  
OS Unidentified.  
XX  
FH Key  
FT CDS  
FT 1..634  
FT /\*tag= a  
FT /product= "Dendritic cell DCLEC/SPLICE 1 variant protein"  
FT /transl\_except= (pos:1, aa:Pro)  
FT /note= "This codon has an apparent deletion of 2  
FT nucleotides, which alters the reading frame; CDS does not  
FT include start codon"  
FT /partial  
XX  
XX WO200172773-A2.  
XX  
XX 04-OCT-2001.  
XX  
XX 28-MAR-2001; 2001WO-EP03542.  
XX  
XX 29-MAR-2000; 2000US-192934P.  
XX  
XX 18-MAY-2000; 2000US-205020P.  
XX  
XX 18-MAY-2000; 2000US-205026P.  
XX  
XX 19-MAY-2000; 2000US-205767P.  
XX  
XX 19-MAY-2000; 2000US-205769P.  
XX  
XX (NOVS ) NOVARTIS AG.  
XX (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.  
XX  
XX Werner G, Phares W, Jaritz M, Lapp H, Kalthoff FS;  
XX  
XX WPI: 2001-616466/71.  
XX P-PSDB; AAE12080.  
XX  
XX New polypeptides for screening therapeutic agonists and antagonists  
XX comprise dendritic cell polypeptides -  
XX  
XX Claim 1; Page 51; 52pp; English.  
XX  
XX The invention relates to dendritic cell (DC) proteins and their  
XX corresponding DNA molecules. A pharmaceutical composition comprising  
XX agonist and antagonist of DC proteins are useful for treating abnormal  
XX conditions related to both an excess of and insufficient level of  
XX expression of DC gene, or related to both an excess of and insufficient  
XX activity of DC protein. Soluble form of DC proteins are used as an active  
XX ingredient in combination with pharmaceutical acceptable carriers.  
XX DC genes and proteins are useful for treating chronic inflammatory  
XX diseases, autoimmune diseases, transplant rejection crisis, including  
XX inflammatory skin diseases such as contact hypersensitivity, atopic  
XX dermatitis or virally-induced immune suppression such as AIDS and cancer.

CC DC protein is useful for inducing immunological response in a mammal, and  
CC as immunogen to produce antibodies immunospecific for the polypeptide.  
CC DC gene is useful in gene therapy. DC gene is also useful as a diagnostic  
CC reagent, and for chromosomal identification. The present sequence is  
CC dendritic cell (DC) DCLEC/SPICE 1 variant gene which is found to belong  
CC to the family of C-type lectins with one single carbohydrate recognition  
CC domain at the C-terminal end.  
XX

SQ Sequence 800 BP; 231 A; 160 C; 183 G; 226 T; 0 other;

Query Match 42.4%; Score 265.6; DB 22; Length 800;  
Best Local Similarity 66.8%; Pred. No. 1e-75;  
Matches 398; Conservative 0; Mismatches 189; Indels 9; Gaps 1;

QY 41 GCTGTTCTCCCTGAGACTCTGTCTCTGTGGGATTTCCATTTGCACCTCTCAGTGGCTT 100  
DB 36 GGTGGTCCAGTTCAGGCTCTGTGCTCCATGCTCATCCATCTTGTCTCTCAGTGGCT 95  
QY 101 CCTTCATTTGAGCTGTAGTACCTTACCATTTTACATATGTTGAAGTGGCAAGGC 160  
DB 96 GTTTCACGTGAGTCTGTGCTGCTCACAATTTTATGATAGCAAACTGTCAAGAGGC 155  
QY 161 TGTCTGAACATA-----CACTCATATCATTTCAAGTCTCACCTGCTTTCAGTGAAGGA 211  
DB 156 TGTCCAGCTTACGAGATATCAACAGTATCATCCAAAGCTGACCTGCGTCATGGAAGGA 215  
QY 212 CAAGGTCGCCAGCTGGGGATGTTGCCAGCTTCTTGGAGTCAATTTGGTTCAGTGGCT 271  
DB 216 AGGACATAGAGATGAGGCTGTGCGCCAAACCTTTGGACTTTCATTTTCAGTCTAGTGGCT 275  
QY 272 ACTTCATTTCCAGTGAAGAGAGGTTTGGTCTTAAGAGTCAGCAGAACTGTGTTGAGATGG 331  
DB 276 ACTTTATTTCTACTGGGATGCAATCTTGGACTTAAGAGTCAAAAGAACTGTTCTGTGATGG 335  
QY 332 GAGCACATTTGGTGTGTTGCACAGACGACAGCAATTTTCATTTGTCAGCAGCTGA 391  
DB 336 GGGCTGTATGTTGGTGTATCAACACCAAGGGAAGCAAGATTTTCATTTTCAGATTTCA 395  
QY 392 ATGAGTCAATTTCTTATTTCTGGGGCTTTTCAGACCCCAAGGTAATTAATTTGGCAAT 451  
DB 396 AAGAATAATTTCTTATTTCTGGGGCTGTACATCCAGGGGTGCGGCACATTTGGCAAT 455  
QY 452 GGATTTGATAGACACCTTATGAGAAATGTCAGATTTTGGCACCTAGTGTAGGCCAATC 511  
DB 456 GGGTTGACAGACACCACTACAATAAGTGAATTTCTGGCACTCAGGTGAACCCATA 515  
QY 512 ATTCCTGACAGCAATGTGCTTCAATAGTCTTCTGGAACCTACAGCATGGGCTGGAATG 571  
DB 516 ACCTTGATGAGGTTGTGGGATAATAATTTCCGTTCTTTCAGAGAATGGGCTGGAATG 575  
QY 572 ATGTTATCTGTAACACTAGAGGAATTCATATGTGAGATGAATAGATTTACCTA 627  
DB 576 ACATTTCACTGTCATGTACCTCAGAGTCAATTTGCAAGATGAAGAGATCTACATA 631

RESULT 11

AAK63162  
ID AAK63162 standard; cDNA; 324 BP.

XX AC AAK63162;

XX DT 06-NOV-2001 (first entry)

XX DE Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:8222.

XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
cytostatic; gene therapy; vaccine; metastasis; ss.

XX OS Homo sapiens.

XX PN WO200157182-A2.

XX XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US01354.  
XX 31-JAN-2000; 2000US-0179065.  
XX 04-FEB-2000; 2000US-0180628.  
XX 24-FEB-2000; 2000US-0184664.  
XX 02-MAR-2000; 2000US-0186350.  
XX 16-MAR-2000; 2000US-0189874.  
XX 17-MAR-2000; 2000US-0190076.  
XX 18-APR-2000; 2000US-0198123.  
XX 19-MAY-2000; 2000US-0205515.  
XX 07-JUN-2000; 2000US-0209467.  
XX 28-JUN-2000; 2000US-0214886.  
XX 30-JUN-2000; 2000US-0215135.  
XX 07-JUL-2000; 2000US-0216647.  
XX 07-JUL-2000; 2000US-0216880.  
XX 11-JUL-2000; 2000US-0217487.  
XX 11-JUL-2000; 2000US-0217496.  
XX 14-JUL-2000; 2000US-0218290.  
XX 26-JUL-2000; 2000US-0220963.  
XX 26-JUL-2000; 2000US-0220964.  
XX 14-AUG-2000; 2000US-0224518.  
XX 14-AUG-2000; 2000US-0224519.  
XX 14-AUG-2000; 2000US-0225213.  
XX 14-AUG-2000; 2000US-0225214.  
XX 14-AUG-2000; 2000US-0225266.  
XX 14-AUG-2000; 2000US-0225267.  
XX 14-AUG-2000; 2000US-0225268.  
XX 14-AUG-2000; 2000US-0225270.  
XX 14-AUG-2000; 2000US-0225447.  
XX 14-AUG-2000; 2000US-0225757.  
XX 14-AUG-2000; 2000US-0225758.  
XX 14-AUG-2000; 2000US-0225759.  
XX 18-AUG-2000; 2000US-0226279.  
XX 22-AUG-2000; 2000US-0226681.  
XX 22-AUG-2000; 2000US-0226868.  
XX 22-AUG-2000; 2000US-0227182.  
XX 23-AUG-2000; 2000US-0227009.  
XX 30-AUG-2000; 2000US-0228924.  
XX 01-SEP-2000; 2000US-0229287.  
XX 01-SEP-2000; 2000US-0229343.  
XX 01-SEP-2000; 2000US-0229344.  
XX 01-SEP-2000; 2000US-0229345.  
XX 05-SEP-2000; 2000US-0229509.  
XX 05-SEP-2000; 2000US-0229513.  
XX 06-SEP-2000; 2000US-0230437.  
XX 06-SEP-2000; 2000US-0230438.  
XX 08-SEP-2000; 2000US-0231242.  
XX 08-SEP-2000; 2000US-0231243.  
XX 08-SEP-2000; 2000US-0231244.  
XX 08-SEP-2000; 2000US-0231413.  
XX 08-SEP-2000; 2000US-0231414.  
XX 08-SEP-2000; 2000US-0232080.  
XX 08-SEP-2000; 2000US-0232081.  
XX 12-SEP-2000; 2000US-0231968.  
XX 14-SEP-2000; 2000US-0232397.  
XX 14-SEP-2000; 2000US-0232398.  
XX 14-SEP-2000; 2000US-0232399.  
XX 14-SEP-2000; 2000US-0232400.  
XX 14-SEP-2000; 2000US-0232401.  
XX 14-SEP-2000; 2000US-0233063.  
XX 14-SEP-2000; 2000US-0233064.  
XX 14-SEP-2000; 2000US-0233065.  
XX 21-SEP-2000; 2000US-0234223.  
XX 21-SEP-2000; 2000US-0234274.  
XX 25-SEP-2000; 2000US-0234997.  
XX 26-SEP-2000; 2000US-0234998.  
XX 27-SEP-2000; 2000US-0235484.  
XX 27-SEP-2000; 2000US-0235834.  
XX 27-SEP-2000; 2000US-0235836.  
XX 29-SEP-2000; 2000US-0236327.  
XX 29-SEP-2000; 2000US-0236367.  
XX 29-SEP-2000; 2000US-0236368.

```
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI: 2001-483426/52.
XX P-PSDB; RAM90381.
XX

XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and
XX metastasis.
XX Claim 1: SEQ ID NO 8222; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
XX amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patients own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting
XX the nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/hematopoietic-related diseases, especially
XX cancers and cancer metastases of hematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/hematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK94950 and AAK82169
XX represent sequences used in the exemplification of the present invention.
XX
XX Sequence 324 BP; 84 A; 72 C; 80 G; 86 T; 2 other;
XX
Query Match 33.0%; Score 207.2; DB 22; Length 324;
Best Local Similarity 91.3%; Pred. No. 6.1e-57;
Matches 251; Conservative 1; Mismatches 20; Indels 3; Gaps 3;
QY 1 ATGATGCAAGAGCAGCAACCTCAAAGTACAGAGAAAGAGCGTGGTTCCTCGAGACTC 60
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 49 ATGATGCAAGAGCAGCAACCTCAAAGTACAGAGAAAGAGCGTGGTTCCTCGAGACTC 108
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 61 TGGTCTGTGGCTGGGATTCATATGCACTCTCAGTCGCTTG-CTTCATTGTCAGCTGTGT 119
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 109 TGGTCTGTGGCTGGGATTCATATGCACTCTCAGTCGCTTG-CTTCATTGTCAGCTGTGT 168
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 120 AGTAACCTTACCATTTCATATGCTGAACTGG-CAAAAGCGTGTCTGAACACACATCAT 178
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 169 AGTAACCTTACCATTTCATATGCTGAACTGG-CAAAAGCGTGTCTGAACACACATCAT 228
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 179 ATCATTCAGTCTCACCTGCTTCAGTGAAGGACAAAGGTCGAGCCT-GGGGATGTTGC 237
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 229 ATCATTCAGTCTCACCTGCTTCAGTGAAGGACAAAGGTCGAGCCT-GGGGATGTTGC 288
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 238 CCAGCTTCTTGGAGTCAATTTGGTTCAGTTGCTA 272
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 289 CCAGCTTCTTGGAGTCAATTTGGTTCAGTTGCTA 323
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 12
ABK52914
ID ABK52914 standard; cDNA; 444 BP.
XX
XX AC ABK52914;
XX
XX 27-AUG-2002 (first entry)
XX
XX Human cDNA encoding a partial dendritic cell immunoreceptor #2.
XX
XX Human; ss; gene; dendritic cell immunoreceptor; cytostatic;
XX antiasthmatic; anorectic; antidiabetic; cancer; allergy; anaphylaxis;
XX asthma; inflammation; obesity; diabetes; central nervous system disorder;
XX Alzheimer's disease; Parkinson's disease; dementia; osteoarthritis;
XX cardiovascular disorder; myocardial infarction; ischaemic heart disease;
XX congestive heart failure; chronic obstructive pulmonary disease; COPD.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..444
```



CC pulmonary disease (COPD) and osteoarthritis (many other diseases and  
CC disorders are listed in the specification). The polypeptide and nucleic  
CC acid are useful for identifying test compounds which act as agonists or  
CC antagonists, for raising specific antibodies, which act as a bait protein in a  
CC two-hybrid or three-hybrid assay. The nucleic acid is useful in  
CC diagnostic assays for detecting diseases and abnormalities or  
CC susceptibility to disease and abnormalities related to the presence of  
CC mutations. The present sequence is a cDNA encoding a partial  
CC dendritic cell immunoreceptor.

XX  
SQ Sequence 402 BP; 114 A; 80 C; 96 G; 112 T; 0 other;

Query Match 32.0%; Score 200.4; DB 24; Length 402;  
Best Local Similarity 68.7%; Pred. No. 1.1e-54;  
Matches 276; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 226 TGGGGATGTTGCCAGCTTCTTGGAAAGTCATTTGGTTCAGTTGCTACTTCAATTTCCAGT 285  
Db 1 TGGAGCTGCTGCCAACCCCTTGGACTTCATTTTCAGTCTAGTTGCTACTTTATTTCTACT 60  
QY 286 GAAGACAAGTTTGGTCTAAGAGTGAGCAGCAACTGTTTGAGATGGGAGCACATTTGGTT 345  
Db 61 GGGATCAATCTTGGACTAAGAGTCAAAAGAACTGTTCTGTGATGGGGCTGATCTGGTG 120  
QY 346 GTCTTCAACACAGACAGCAGAGAAATTTTCATTTGTCAGCAGCTGAATGAGTCATTTTCT 405  
Db 121 GTATCAACACAGGAGAAACAGGATTTTCATTCAGATCTGAAGAAATTTCTTCT 180  
QY 406 TATTTTCTGGGGCTTTTCAGACCCACAGGTAATAATAATTTGCAATGGATTTGATGAGACA 465  
Db 181 TATTTTCTGGGGCTGTTCAGATCCAGGGGTGCGCGACATTTGCAATGGTTCACCGACACA 240  
QY 466 CTTATGAGAAAATGTCAGATTTGCGACCTAGGTAGGCGCAATTCATCTTCAGAGCAA 525  
Db 241 CCATCAATGAAATGTCAGATTTGCGACTTCTGGCACTCAGGTGAAACCAATACCTTGATGAGCT 300  
QY 526 TGTGCTTCAATAGTCTTCTGGAACCTACAGGATGGGGCTGGAATGATGATTTATCTGTGAA 585  
Db 301 TGTGCGATAATAAATTTCCGTTCTTCAGAAGAATGGGCTGGAATGATGATTCACCTGTCAT 360  
QY 586 ACTAGAAGGAATCAATATGTGAGATGAATAAGATTTACCTA 627  
Db 361 GTACCTCAGAAGTCAATTTTCAAGATGAAGAAGATCTACATA 402

RESULT 14

AAS31385  
ID AAS31385 standard; cDNA; 1091 BP.

XX  
AC AAS31385;

DT 04-DEC-2001 (first entry)

XX Human cDNA encoding a novel extracellular matrix protein, Seq ID No 199.

XX Human; secreted extracellular matrix protein; ss; immunomodulatory;  
XX Anti-HIV; antianemic; antirheumatic; antisclerotic; cardiac; vascular;  
XX cerebroprotective; thrombolytic; antimicrobial; ophthalmic; cytostatic;  
XX antialzheimers; immune/autoimmune disease; HIV infection; anaemia;  
XX human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;  
XX cancers; hyperproliferative disorder; breast neoplasm; melanoma;  
XX Sazary syndrome; Gaucher's disease; neurological diseases;  
XX Alzheimer's disease; Parkinson's disease; cardiovascular disorder;  
XX cardiac arrest; tachycardia; angina; infection; corneal infections;  
XX wound healing; immunogen; gene therapy; antisense; food additive.

XX Homo sapiens.

OS WO20015368-A1.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01348.

XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.



PR	02-OCT-2000;	2000US-0236802;
PR	02-OCT-2000;	2000US-0237037.
PR	02-OCT-2000;	2000US-0237038.
PR	02-OCT-2000;	2000US-0237039.
PR	02-OCT-2000;	2000US-0237040.
PR	13-OCT-2000;	2000US-0239935.
PR	13-OCT-2000;	2000US-0239937.
PR	20-OCT-2000;	2000US-0240960.
PR	20-OCT-2000;	2000US-0241221.
PR	20-OCT-2000;	2000US-0241785.
PR	20-OCT-2000;	2000US-0241786.
PR	20-OCT-2000;	2000US-0241787.
PR	20-OCT-2000;	2000US-0241808.
PR	20-OCT-2000;	2000US-0241809.
PR	20-OCT-2000;	2000US-0241826.
PR	01-NOV-2000;	2000US-0244617.
PR	01-NOV-2000;	2000US-0244618.
PR	08-NOV-2000;	2000US-0246474.
PR	08-NOV-2000;	2000US-0246475.
PR	08-NOV-2000;	2000US-0246476.
PR	08-NOV-2000;	2000US-0246477.
PR	08-NOV-2000;	2000US-0246478.
PR	08-NOV-2000;	2000US-0246523.
PR	08-NOV-2000;	2000US-0246524.
PR	08-NOV-2000;	2000US-0246525.
PR	08-NOV-2000;	2000US-0246526.
PR	08-NOV-2000;	2000US-0246527.
PR	08-NOV-2000;	2000US-0246528.
PR	08-NOV-2000;	2000US-0249207.
PR	17-NOV-2000;	2000US-0249208.
PR	17-NOV-2000;	2000US-0249209.
PR	17-NOV-2000;	2000US-0249210.
PR	17-NOV-2000;	2000US-0249211.
PR	17-NOV-2000;	2000US-0249212.
PR	17-NOV-2000;	2000US-0249213.
PR	17-NOV-2000;	2000US-0249214.
PR	17-NOV-2000;	2000US-0249215.
PR	17-NOV-2000;	2000US-0249216.
PR	17-NOV-2000;	2000US-0249217.
PR	17-NOV-2000;	2000US-0249218.
PR	17-NOV-2000;	2000US-0249244.
PR	17-NOV-2000;	2000US-0249245.
PR	17-NOV-2000;	2000US-0249264.
PR	17-NOV-2000;	2000US-0249265.
PR	17-NOV-2000;	2000US-0249267.
PR	17-NOV-2000;	2000US-0249269.
PR	01-DEC-2000;	2000US-0250160.
PR	01-DEC-2000;	2000US-0250391.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	06-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	06-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251990.
PR	11-DEC-2000;	2000US-0254097.
XX	05-JAN-2001;	2001US-0259678.
PA	(HUMA-) HUMAN GENOME SCI INC	
PI	Rosen CA, Barash SC, Ruben	
XX	WPI: 2001-465572/50.	
DR	P-PSDB: AAU19814.	
XX	Nucleic acid molecules encod	
PT		

proteins, used in preventing, treating or ameliorating a disorder, e.g. Alzheimer's and Parkinson's diseases and cancers -

Claim 1; SEQ ID No 199; 577pp; English.

The invention relates to isolated nucleic acid molecules encoding novel human secreted extracellular matrix proteins (SPs). The polynucleotides and proteins are used to prevent, treat a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. For example, disorders associated with decreased expression of SPs. The SP polynucleotide or a vector expressing them may be administered to treat diseases by gene therapy. Antisense molecules may be administered to down regulate expression of SPs by binding with the cells own genes and preventing their expression. The polynucleotides may also be used as DNA probes in diagnostic assays. The SPs may also be used as antigens to produce antibodies and to identify modulators (agonists and antagonists) of the SPs. The anti-(SP) antibodies and antagonists may also be used to down regulate expression and activity of SP and as diagnostic agents for detecting the presence of SPs in samples. The disorders include for example: immune/autoimmune diseases (e.g. HIV (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis and multiple sclerosis), cancers and hyperproliferative disorders (e.g. melanomas, neoplasms of the breast or liver, Sezary syndrome and Gaucher's disease), neurological diseases (e.g. Alzheimer's disease, Parkinson's disease) cardio-/cerebrovascular disorders (e.g. cardiac arrest, tachycardia and angina), infections caused by bacteria, viruses and fungi and ocular disorders (e.g. corneal infections). Other uses include wound healing, maintenance of organs before transplantation,

Query Match 30.6%; Score 192; DB 22; Length 1091;  
Best Local Similarity 68.0%; Pred.No. 1e-51;  
Matches 283; Conservative 0; Mismatches 130; Indels 3; Gaps 1

Qy 215 AGGTGCCAGCGTGGGATGTGCCAGCTTC TTGGAAGTCATTGGTTCCAGTTGCTACT 274  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 314 AAGACAGCGCTGGAGCTGTGCCAAACAATTGGAAGTCATTAGTTCCAAGTCTACT 373  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Qy 275 TCATTTCCAGTGGAAGAAGGTTTGCTTAAGAGTGAGCAGCACTGTGTCAGATGGGAG 334  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 374 TTAFTTCTACTGAATCAGCATCTTGCAAGACAGTGAGAAGCACTGCTAGAATGGAGG 433  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Qy 335 CACATTTGGTTGTTCACACAGACGACAGCAAGATTTCAATGTCACAGCAGCTGAATG 394  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 434 CTCACCTGCGTGTGATAAACACTCAAGAAGACGAGCATTTTCATCTCCAGAATCTGCAAG 493  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Qy 395 AGTCATTTCTTATTTTCTGGGCTTTCAGACCACAGGTAATAAATTTGCAATGGA 454  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 494 AAGAATCTGCTTATTTTGTGGGCTCTCAGATCCAGAGGTGACGCACATTTGGCAANTGGG 553  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Qy 455 TTGATAAGACACCTTATGAAAAAATGTCCAGATTTTGGCACCTAGGTGAGGCCCAATCAT 514  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 554 TTGATCAGACACCATCAATGAAGATTCACATTTCTGGCATCCAGTCAGCCCCAGTGATC 613  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Qy 515 CTGCAGACGAATGTGCTTCATAGTCTTCTGGAA ---ACCTCAGGATGGGCTGGGAATG 571  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 614 CCAATGACGCGTGGTTGCTAAATTTTCGTAAATTCACCCAAAAGTGGGCTGGGAATG 673  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Qy 572 ATGTTATCTGTGAACCTAGAGAATTCATATGTGAGATGAATAAGATTTTACCTA 627  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 674 ATGTTAATGTCTGTGGTCTCAAGGTCAAGTCTGTTGTGAGATGATGAAGATTCACCTTA 729  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 15  
ID AB066709 standard; cDNA; 1091 BP.  
XX AB066709;  
AC AB066709;  
DT 23-AUG-2002 (first entry)  
XX Human polynucleotide SEQ ID NO 199.  
DE Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;  
KW



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 17, 2003, 08:49:16 : Search time 22.9615 Seconds  
(without alignments)  
13907.882 Million cell updates/sec

Title: US-09-766-511B-52

Perfect score: 627

Sequence: 1 atgatgaagagcagcaacc.....agatgaataagatttaccta 627

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 424239 seqs, 254661826 residues

Total number of hits satisfying chosen parameters: 848478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published\_Applications\_NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
c 1	355.6	56.7	693	10	US-09-833-381-994
c 2	311.6	49.7	528	10	US-09-833-381-995
c 3	273.4	43.6	642	12	US-10-090-466-1
4	231	36.8	549	12	US-10-090-466-3
5	192	30.6	1091	10	US-09-764-870-199
6	192	30.6	1096	10	US-09-764-870-38
7	192	30.6	1104	9	US-09-862-802-1
c 8	181.8	29.0	1036	10	US-09-833-381-119
c 9	151.2	24.1	291	10	US-09-833-381-996
10	134	21.4	1418	9	US-09-862-802-7
11	120.6	19.2	758	10	US-09-833-381-1339
12	117.4	18.7	968	9	US-09-965-529-40
13	117.4	18.7	997	9	US-09-905-291A-376
14	117.4	18.7	997	9	US-09-902-853-376
15	117.4	18.7	997	9	US-09-907-824-376
16	117.4	18.7	997	9	US-09-907-841-376
17	117.4	18.7	997	9	US-09-904-011-376
18	117.4	18.7	997	9	US-10-174-590-23
19	117.4	18.7	997	9	US-10-176-758-23

20	117.4	18.7	997	9	US-10-175-737-23	Sequence 23, Appl
21	117.4	18.7	997	9	US-09-906-742-376	Sequence 376, App
22	117.4	18.7	997	9	US-10-173-706-23	Sequence 23, Appl
23	117.4	18.7	997	9	US-10-175-738-23	Sequence 23, Appl
24	117.4	18.7	997	9	US-10-175-752-23	Sequence 23, Appl
25	117.4	18.7	997	9	US-10-176-482-23	Sequence 23, Appl
26	117.4	18.7	997	9	US-10-176-757-23	Sequence 23, Appl
27	117.4	18.7	997	9	US-10-176-913-23	Sequence 23, Appl
28	117.4	18.7	997	9	US-10-180-552-23	Sequence 23, Appl
29	117.4	18.7	997	9	US-10-180-557-23	Sequence 23, Appl
30	117.4	18.7	997	9	US-09-906-838-376	Sequence 376, App
31	117.4	18.7	997	9	US-09-907-613-376	Sequence 376, App
32	117.4	18.7	997	9	US-09-907-942-376	Sequence 376, App
33	117.4	18.7	997	9	US-10-173-700-23	Sequence 23, Appl
34	117.4	18.7	997	9	US-10-174-572-23	Sequence 23, Appl
35	117.4	18.7	997	9	US-10-174-579-23	Sequence 23, Appl
36	117.4	18.7	997	9	US-10-174-582-23	Sequence 23, Appl
37	117.4	18.7	997	9	US-10-174-588-23	Sequence 23, Appl
38	117.4	18.7	997	9	US-10-175-739-23	Sequence 23, Appl
39	117.4	18.7	997	9	US-10-175-740-23	Sequence 23, Appl
40	117.4	18.7	997	9	US-10-175-743-23	Sequence 23, Appl
41	117.4	18.7	997	9	US-10-176-488-23	Sequence 23, Appl
42	117.4	18.7	997	9	US-10-176-492-23	Sequence 23, Appl
43	117.4	18.7	997	9	US-10-176-747-23	Sequence 23, Appl
44	117.4	18.7	997	9	US-10-176-750-23	Sequence 23, Appl
45	117.4	18.7	997	9	US-10-176-985-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1

US-09-833-381-994/c

; Sequence 994, Application US/09833381

; Patent No. US20020132090A1

; GENERAL INFORMATION:

; APPLICANT: Robison, Keith E.

; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs

; FILE REFERENCE: 5800-119

; CURRENT APPLICATION NUMBER: US/09/833,381

; CURRENT FILING DATE: 2001-04-11

; PRIOR APPLICATION NUMBER: 09/516,448

; PRIOR FILING DATE: 2000-02-29

; NUMBER OF SEQ ID NOS: 2050

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 994

; LENGTH: 693

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (1)...(693)

; OTHER INFORMATION: n = A,T,C or G

; US-09-833-381-994

Query Match 56.7%; Score 355.6; DB 10; Length 693;

Best Local Similarity 76.5%; Pred. No. 1.6e-99;

Matches 463; Conservative 0; Mismatches 135; Indels 7; Gaps 2;

QY	30	AGAGAAAGAGGCTGCTTCCTCCCTGAGACTCTGGTCTGGTGGGATTTCCATTGCACT	89
Db	690	AGGAGAGGAGTCTGCTGACCCCTGAGACTCTGGTCAGCTGCTGATTTCCATTGCTTACT	631
QY	90	CCTCAGTGTCTGCTCAATTTGAGCTGTGTAGTAACCTTACCATTATTTACATATGGTGAAC	149
Db	630	CTTGACTACCTGTTTCATTTGCGAGCTGTGTGGTGACTTTACCAATTTATTTATGGACGAC	571
QY	150	TGCGAAAGGCTCTCTGAAGTACACTCATATCATTCACAGTCTCACCTGCTTCAGTGAAG	209
Db	570	CAGTAGAAGACATATATGAACCTTCACACATACCATTTCCAGTCTCACCTGCTTCAGTGAAG	511
QY	210	GACAAAGTGGCAG-----CCTGGGGATGTTGCCAGCTTCTTGAAGTCAATTTGGTTC	263





\_\_\_\_\_

433 GGTAAATAATTGGCAATGGATTGATAAGACACACCTTATGAGAAAAATGTC-----

Query Match	29.0%;	Score 181.8;	DB 10;	Length 1036;
Best Local Similarity	65.0%;	Pred. No. 5.6e-46;		
Matches 293. Conservative	0;	Mismatches 142;	Indels 16;	Gaps 1

```
Db 371 GGTGGCGACATTTGGCAATGGTTTGACGAGACACCACATACATAATGAAATGTCAAGTGAGTA 312
Qy 484 -----AGATTTGGCACCTAGGTGAGCCCAATCATTTCTGCAGAGCAATGTCTTCAAT 536
Db 311 TAGAATGAGATTCGGCACTAGCTAGCAACCAATAACCTTGATGAGCGTTGTGCGATAAT 252
Qy 537 AGTCTTCTGGAACACTACAGATGGGCTGGAATGATGTTATCTGTGAAACTAGAAGAA 596
Db 251 AAATTTCCGTCTTTCAGAAGAATGGGCTGGAATGACATTCACCTGTCATCTCAGAA 192
Qy 597 TTCAATATGTCAGATGAATAGATTTACCTA 627
Db 191 GTCAATTTGCAAGATGAAGAGATCTACATA 161

RESULT 9
US-09-833-381-996/c
; Sequence 996, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 996
; LENGTH: 291
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(291)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-996

Query Match 24.1%; Score 151.2; DB 10; Length 291;
Best Local Similarity 72.5%; Pred. No. 7.5e-37;
Matches 211; Conservative 0; Mismatches 78; Indels 2; Gaps 2;

Qy 264 CAGTGTCTACTTTCATTTCCAGTGAAG-AGAAAGGTTTGGTCTAAGAGTGAGCAGCAACTGTG 322
Db 291 CAGCTGTACTCTCATTTCTACCAGGAAGAACTTCTGNAGCACCAGTGAGCANAAGTNG 232
Qy 323 TTGAGATGGGAGCACATTTGGTTGTGTCA-ACACAGAAGCAGAGCAAGATTTTCATTTGC 381
Db 231 TTCATATGGGGGCTCATCTGNGGTTGATCACATACTCAAGCGGAGCAGANTTTTCATCAC 172
Qy 382 CAGCAGCTGAATGAGTCATTTCTTATTTCTGGGGCTTTCAGACCCACAGGTAATAAT 441
Db 171 CAGCAGCTGAATGAGTCATTTTACTTCTCGGGTCTTTTCGGATCCACAAAGTNNNGG 112
Qy 442 AATTGGCAATGGATTGATAAGACACCTTATGAGAAAATATGTCAGATTTTGGCAGCTAGT 501
Db 111 AAANGNAATGGATGATGATACCTTTCAGTCAAAATGTCAGGTTNTGCGNCCCAT 52
Qy 502 GAGGCCAATCATTTCTGCAGACCAATGTGCTTCAATAGTCTTCTTGGAAACCT 552
Db 51 GAACCCCAATCTCCAGAGAAGNGGTGNCNTCAATAGTTTACTTGAATCCT 1

RESULT 10
US-09-862-802-7
; Sequence 7, Application US/09862802
; Patent No. US20020165346A1
; GENERAL INFORMATION:
; APPLICANT: Schering-Plough Corporation
; TITLE OF INVENTION: MAMMALIAN MEMBRANE PROTEIN GENES; RELATED REAGENTS
; FILE REFERENCE: SF0695B
; CURRENT APPLICATION NUMBER: US/09/862,802
```

```
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/111,470
; PRIOR FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 1418
; TYPE: DNA
; ORGANISM: unknown
; FEATURE:
; OTHER INFORMATION: mammalian nucleic acid
; NAME/KEY: CDS
; LOCATION: 279...992
; OTHER INFORMATION: protein coding sequence
; NAME/KEY: misc_feature
; LOCATION: 1348
; OTHER INFORMATION: poly A addition motif
US-09-862-802-7

Query Match 21.4%; Score 134; DB 9; Length 1418;
Best Local Similarity 61.4%; Pred. No. 3.3e-31;
Matches 253; Conservative 0; Mismatches 150; Indels 9; Gaps 2;

Qy 222 AGCTGGGAGTGTGCCAGCTTCTTTGGAAGTCATTTGGTGTCCAGTTGCTACTTTCATTTTC 281
Db 584 AGTCTGGAGCTGTGTGCCCAAGAGATTGGAGGCTATTTTGGTTCCCACTGCTACTTGGTTCC 643
Qy 282 CAGTGA-----AGAGAAGTTTGGTCTAAGAGTGAGCAGAACTGTGTTGAGATGGGAGC 335
Db 644 CACAGTTTCTTCATCAGCATCTTGGAAACAAGAGTAGGAGAACTGCTCCCGCATGGGTGC 703
Qy 336 ACATTTGGTGTGTTCACACAGACGAGCAGAGATTTTCATTTGTCCACAGCTGAATCA 395
Db 704 TCATCTAGTGGTGATCCAAAGCCAGGAAGCAGAGGATTTTCATCAGCTGGGATCTTGGACAC 763
Qy 396 GTCATTTTCTTATTTTCTGGGGCTTTCAGACCCCAAGGTAATAATAATTGGCAATGGAT 455
Db 764 TCATGCTGCTTATTTTATAGGTTGTGGGATAC---AGGCCATCGGCAATGGCAATGGGT 820
Qy 456 TGATAGACACCTTATGAGAAAAATGTCAGATTTTGGCACCTAGGTGAGCCCAATCATTC 515
Db 821 TGATAGACACCACTATGAAGAAGTATCACATTTCTGGCAACAATGGTGAAGCCAGCAGTGG 880
Qy 516 TGCAGAGCAATGTGCTTCAATAGTCTTCTGGAACCTCACAGGATGGGCTGGAATGATGT 575
Db 881 CAATGAAAATGTGCTACAAATAATTTACCGTTGGAAGACTGGATGGGCTGGAACGATAT 940
Qy 576 TATCTCTGAACTAGAGGAATTCAAATATGTGAGATGAATAAGATTTTACCTA 627
Db 941 CTCTTCAGTCTTTAAACAGAAAGTCAGTTTGTGAGATGAAGAAAATAAACTTA 992

RESULT 11
US-09-833-381-1339
; Sequence 1339, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1339
; LENGTH: 758
; TYPE: DNA
; ORGANISM: Homo sapiens
; OTHER INFORMATION:
US-09-833-381-1339

Query Match 19.2%; Score 120.6; DB 10; Length 758;
```

Best Local Similarity 52.5%; Pred. No. 3.2e-27;  
Matches 321; Conservative 0; Mismatches 274; Indels 16; Gaps 2;  
QY 17 AACCTCAAGTACAGAGAAAAGAGC-TGGTTGTCCCTGAGACTCTGGTCTGTGGCTGGG 75  
Db 18 AACCTCAAGTAAACTGAAGAGGATGCATCCAGCTGCTCGGTTATTGCTGTA 77  
QY 76 ATTTCCATTGGACCTCCTCAGTCTGTGTTCAATTGTGAGCTGTGTAGTAACCTTACCATT 135  
Db 78 GTTTTCATCTTACTCTCAGTGTCTGTTTATTGCAAGTTGTTTGGTGACCTACACAAC 137  
QY 136 ACATATGTGAACCTGCAAAAGGCTGCTGAACACTACACTATATCATTCAGTCTCACC 195  
Db 138 TTTTCACGCTGTAAGAGAGCACAGGAGTGCAAGTTAGAGCACCATTGCAAGCTCAAA 197  
QY 196 TGCTTCAGT-----AAGGGACAAAGGTGCCAGCTCGGGATGTTGCCCA 240  
Db 198 TGCATCAAGAGAAATCAGAACTTAAATACATTAAAGGAGCACCTGGAACGTGTGCT 257  
QY 241 GCTTCTTGGAAAGTCATTTGGTTCCAGTTGCTACTTCTATTTCCAGTGAAGAGAGTTGG 300  
Db 258 ATTGACTGGAGCGCTTCCAGTCCAACTGCTATTCTCTTACTGACAAAGACGTGG 317  
QY 301 TCTAGAGTGGACAGAACTGTTGAGATGGGACACATTTGGTTGTTCAACACAGAA 360  
Db 318 GCTGAGAGTGAAGGAAGCTGTTCAAGGGATGGGGCCCATCTGATGACCATCAGACGGAA 377  
QY 361 CGACAGCAGAAATTTCAATGTCACAGCAGCTGAATGAGTCATTTTCTTATTTCTGGGCTT 420  
Db 378 GCTGAGCAGAACTTATTATTTCAGTTTCTGTGATAGACGGCTTCTCTATTTCTTGGACTT 437  
QY 421 TCAGACCACCAAGGTAATAAATGGCAATGGATGATGAAGACACCTTATGAGAAATAAT 480  
Db 438 AGAGATGAGAATGCCAAAGTCAGTGGGTTGGGTGGACCAAGCCATTTAAACCCACGC 497  
QY 481 GTCAGATTTTGGCACTAGGTGAGCCCAATCATTTCTGCAGAGCAATGTCTTCAATAGTC 540  
Db 498 AGAGATTTCTGGCATTAAGATTAAGCCCGACAACTCTCAGGGAGAAACTGTGTTGTTCTT 557  
QY 541 TTCTGGAACCTTACAGGATGGGCTGGAATGATGTATCTGAAACTAGAAAGAAATTC 600  
Db 558 GTTTTATACCAAGATAAATGGCCCTGGAATGATGTCTCTGTAACCTTGAAGCAAGTAGG 617  
QY 601 ATATCTGAGAT 611  
Db 618 ATTTGTAATAAT 628

RESULT 12

US-09-965-529-40

; Sequence 40, Application US/09965529

; Publication No. US20020182671A1

; GENERAL INFORMATION:

; APPLICANT: LAL, Preeti

; APPLICANT: YUE, Henry

; APPLICANT: TANG, Y. Tom

; APPLICANT: BANDMAN, Olga

; APPLICANT: BURFORD, Neil

; APPLICANT: AZIMZAI, Yalda

; APPLICANT: BAUGHN, Mariah R.

; APPLICANT: LU, Dyrung Aina M.

; APPLICANT: PATTERSON, Chandra

; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS

; FILE REFERENCE: PF-0731 USA

; CURRENT APPLICATION NUMBER: US/09/965,529

; PRIOR FILING DATE: 2001-09-26

; PRIOR APPLICATION NUMBER: 60/149,641; 60/164,203; PCT/US00/22315

; PRIOR FILING DATE: 1999-08-17; 1999-11-09; 2000-08-14

; NUMBER OF SEQ ID NOS: 74

; SOFTWARE: PERL Program

; SEQ ID NO 40

; LENGTH: 968

; TYPE: DNA

; ORCANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No. US20020182671A1 1521513CB1  
US-09-965-529-40

Query Match 18.7%; Score 117.4; DB 9; Length 968;

Best Local Similarity 54.4%; Pred. No. 3.5e-26;

Matches 329; Conservative 0; Mismatches 261; Indels 15; Gaps 4;

QY 17 AACCTCAAGTACAGAGAAAAGAGCTGTGTTGCC--TGAGACTCTGGTCTGTGGCTG 73

Db 184 AACACAAATGCACAGAGAGAGGATGCTTCTCTCCCAATGTCTTATGGACTGTGCTG 243

QY 74 GGATTTCATTGCACCTCCCTCAGTGTGCTTCATTGTGTAGCTGTGTAGTAACCTTACCATT 133

Db 244 GGATCCCATCTCTATTCTCAGTGCCTGTTTCATCACCAGATGTGTGTGACATTTCCGA 303

QY 134 TTACATATGTTGAACTGGCAAAAGCTGTCTGAACCTACACTCATATCATTCAAAGTCTCA 193

Db 304 TCTTTCAAAACCTGTGTATGAGAAAAGTT---TCAGCTACCTGAGAAATTTTCACAGAGCTCT 360

QY 194 CTTGCTTTCAGTGAAGGGACAAAGTCCAGCTGGGGATGTTGCCAGCTTCTTGGAAAGT 253

Db 361 CTTGCTACAAATATGATCA---GGTTCAGTCAAGAAATTTGTTCCATTGAACTGGGAAT 417

QY 254 CATTGGTTCCAGTTGCTACTTCTCATTTCCAGTGAAGAGAGAGTTGGTCTAAGAGTGAGC 313

Db 418 ATTTTCAATCCAGCTCTACTTCTTTCTTACTGACACCAATTTCTCTGGGGTTTAAAGTTTAA 477

QY 314 AGNACTGTGTGAGATGGGAGCAGACATTTGGTTGTGTTCAACACAGACAGAGCAGAAAT 373

Db 478 AGAAGTCTCACCAATGGGGGCTCACTGGTGTATCAACTCACAGGAGGAGCAGGAAT 537

QY 374 TCATTCTCAGCAGCTGAATGAGTCATTTTCTTATTCTCTGGGCTTTTCAGACCCACACAAG 433

Db 538 TCCTTTCTTACAAGAAACCTAAATGAGAGAGTTTTATTGAGCTGTGACAGCAGGTTG 597

QY 434 GTAATAAATATGGCAATGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 493

Db 598 TCGAGGCTCAGTGGCAATGGTGGAGCGCACACCTTTGACAAAAGTCTCTGAGCTTCTGGG 657

QY 494 ACCTAGTGGAGCCCAATCA-----TTCTGACAGACAATGCTTCAATAGTCTTCTGGA 547

Db 658 ATGTAGGGAGCCCCAACACATAGCTACCTGTGAGGAGCTGTGCCACCATGAGAGACTCTT 717

QY 548 AACCTACAGGATGGGCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 607

Db 718 CAAACCCAGGCCAAATTTGGAATGATGATGATGATGATGATGATGATGATGATGATGATG 777

QY 608 AGATG 612

Db 778 AAATG 782

RESULT 13

US-09-905-291A-376

; Sequence 376, Application US/09905291A

; Patent No. US20020160374A1

; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.

; APPLICANT: Ashkenazi, Avi

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Fong, Sherman

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, A.

; APPLICANT: Godowski, Paul J.





```

; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 376
; LENGTH: 997
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-902-853-376

```

Query Match 18.7%; Score 117.4; DB 9; Length 997;  
Best Local Similarity 54.4%; Pred. NO. 3.6e-26;  
Matches 329; Conservative 0; Mismatches 261; Indels 15; Gaps 4;

Qy	17	AACTC	AAAGT	CTGGT	GTGCC	---	TGAC	ACTCT	GGT	CTGTGG	CGTG	73
Db	128	AAACAA	TGACAGAGAGAGG	ATGCTCT	TCTCC	AAATGTTCT	TATG	AGACTG	TGTTG	TCG	187	
Qy	74	GGATT	TCATTCAGT	CTCCTCAG	TGCTTGTCTT	CAATG	TGAG	CTGTG	TAGTAA	CTTAC	133	
Db	188	GGAT	TCCCATCTT	ATTTCTCAG	TGCGCTGTTT	CA	TAC	CACG	ATGTTG	TGAC	247	
Qy	134	TTAC	ATATG	TGAAC	TGGCA	AAAGCTG	TC	TGA	ACTCAT	ATAT	193	
Db	248	TC	TTTCA	AACTCTG	TATG	TAGAAA	AGTT--	-TC	AG	CTACTG	AGA	304
Qy	194	CTG	CTTCA	G	TGA	GGGAC	AAAGT	GCC	AGC	CTGG	GATG	253
Db	305	CCTG	CTACA	ATAT	ATGG	ATCA	---	GGT	CTA	AA	GAATTT	361
Qy	254	CATT	TGGT	CCAG	TGCT	ACTTC	ATT	CC	AGTGA	AGAG	GTTT	313
Db	362	ATT	TTCA	ATCC	AGT	GCTACT	CTT	TTT	CTACT	GAC	ACCAT	421
Qy	314	AGAA	CTCTG	TGAT	GGGAG	CAC	ATT	TGG	TGTG	TTC	AAC	373
Db	422	AGAA	CTCTC	AGC	CAATGGG	GGCTCA	CGT	TGG	TGTAT	CA	AACTC	481
Qy	374	TCAT	TG	CCAG	CAGTGA	ATG	ATCTTT	CT	TATTT	CT	TGG	433
Db	482	TC	CTTT	CCT	CA	AGAA	ACCT	TA	AAAT	TG	AG	541
Qy	434	GTAA	TAA	TAT	TGCA	ATG	GAT	TG	ATA	AG	ACCT	493
Db	542	TC	GAGG	TCAGT	GGCA	TG	GTGG	AGG	CAC	ACCT	TG	601

QY	494	ACCTAGCTGAGCCCAATCA-----TCTCGACAGCAATGTGCTCTCAATAGTCTTCTGTGGA	544
Db	602	ATGTAGGGGAGCCCAACAACATAGTACCCCTGGAGGACTGTGGCCACCATGAGAGACTCTT	661
QY	548	AACCTACAGGATGGGCTGGAAATCATCTTATCTGTGTAAGAACTAGAAGAAATTCAAATATGTG	607
Db	662	CAMACCAAGGCAAAATTTGGAATGATTAACCTGTTTCTCTCAATTTATTTTCGAGATTGTG	721
QY	608	AGATG 612	
Db	722	AAATG 726	
RESULT 15			
US-09-907-824-376			
; Sequence 376, Application US/09907824			
; Publication No. US20020197671A1			
; GENERAL INFORMATION:			
; APPLICANT: Genentech, Inc.			
; APPLICANT: Ashkenazi, Avi			
; APPLICANT: Botstein, David			
; APPLICANT: Desnoyers, Luc			
; APPLICANT: Eaton, Dan L.			
; APPLICANT: Ferrara, Napoleone			
; APPLICANT: Filvaroff, Ellen			
; APPLICANT: Fong, Sherman			
; APPLICANT: Gao, Wei-Qiang			
; APPLICANT: Gerber, Hanspeter			
; APPLICANT: Gerriksen, Mary E.			
; APPLICANT: Goddard, A.			
; APPLICANT: Godowski, Paul J.			
; APPLICANT: Grimaldi, Christopher J.			
; APPLICANT: Gurney, Austin L.			
; APPLICANT: Hillan, Kenneth, J.			
; APPLICANT: Klijavin, Ivar J.			
; APPLICANT: Mather, Jennie P.			
; APPLICANT: Pan, James			
; APPLICANT: Paoni, Nicholas F.			
; APPLICANT: Roy, Margaret Ann			
; APPLICANT: Stewart, Timothy A.			
; APPLICANT: Tumas, Daniel			
; APPLICANT: Williams, P. Mickey			
; APPLICANT: Wood, William, I.			
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic			
; TITLE OF INVENTION: Acids Encoding the Same			
; FILE REFERENCE: 10466-14			
; CURRENT APPLICATION NUMBER: US/09/907,824			
; CURRENT FILING DATE: 2001-07-17			
; PRIOR APPLICATION NUMBER: 09/665,350			
; PRIOR FILING DATE: 2000-09-18			
; PRIOR APPLICATION NUMBER: PCT/US00/04414			
; PRIOR FILING DATE: 2000-02-22			
; PRIOR APPLICATION NUMBER: US 60/143,048			
; PRIOR FILING DATE: 1999-07-07			
; PRIOR APPLICATION NUMBER: US 60/145,698			
; PRIOR FILING DATE: 1999-07-26			
; PRIOR APPLICATION NUMBER: US 60/146,222			
; PRIOR FILING DATE: 1999-07-28			
; PRIOR APPLICATION NUMBER: PCT/US99/20594			
; PRIOR FILING DATE: 1999-09-08			
; PRIOR APPLICATION NUMBER: PCT/US99/20944			
; PRIOR FILING DATE: 1999-09-13			
; PRIOR APPLICATION NUMBER: PCT/US99/21090			
; PRIOR FILING DATE: 1999-09-15			
; PRIOR APPLICATION NUMBER: PCT/US99/21547			
; PRIOR FILING DATE: 1999-09-15			
; PRIOR APPLICATION NUMBER: PCT/US99/23089			
; PRIOR FILING DATE: 1999-10-05			
; PRIOR APPLICATION NUMBER: PCT/US99/28214			
; PRIOR FILING DATE: 1999-11-29			
; PRIOR APPLICATION NUMBER: PCT/US99/28313			
; PRIOR FILING DATE: 1999-11-30			
; PRIOR APPLICATION NUMBER: PCT/US99/28564			

Search completed: February 17, 2003, 11:51:36  
Job time : 26.9615 secs



GenCore version 5.1.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 17, 2003, 08:47:16 ; Search time 602.362 Seconds  
(without alignments)  
16857.906 Million cell updates/sec

Title: US-09-766-511B-52  
Perfect score: 627  
Sequence: 1 atgatgaagagcagcaacc.....agatgaataagatttaccta 627

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	332.4	53.0	504	12	BF603486
2	320	51.0	470	12	BF774158
3	310	49.4	971	13	BI696642
c 4	176	28.1	582	9	AI183967
c 5	172.8	27.6	753	9	AI949910
6	167.2	26.7	715	10	AV716802

c	7	166.8	26.6	1013	9	AA868502
	8	166	26.5	582	10	AV716725
	9	151.4	24.1	514	12	BF152383
	10	151.4	24.1	558	9	AI391056
	11	141.8	22.6	561	10	BE650539
	12	140	22.3	631	11	AK007794
	13	139.4	22.2	1200	11	AK014135
	14	135	21.5	448	9	AA446401
	15	135	21.5	666	10	AW242656
	16	134.6	21.5	812	11	AK020363
	17	134	21.4	568	10	BE200183
	18	130.6	20.8	1007	14	BQ715169
	19	130.2	20.8	403	10	AW822082
	20	119.8	19.1	498	10	BE689245
	21	116.8	18.6	889	12	BG164873
	22	116.4	18.6	1032	10	BE311916
	23	108.6	17.3	608	14	BQ002993
	24	106.4	17.0	667	10	BE688920
	25	105.6	16.8	681	12	BF020383
	26	104.6	16.7	334	10	BE690187
	27	103.4	16.5	831	12	BG065735
	28	99.8	15.9	678	12	BF227688
	29	96.6	15.4	559	10	BE120531
	30	94.8	15.1	576	12	BG079007
	31	94.4	15.1	676	13	BI909298
	32	92.6	14.8	738	14	BQ004812
	33	90.4	14.4	408	10	AV711676
	34	89.4	14.3	544	9	A1400631
	35	88.8	14.2	264	9	AV057557
	36	87.4	13.9	432	10	BB863598
	37	87.2	13.9	228	9	AV252043
	38	86.8	13.8	458	10	BE632986
	39	85.8	13.7	291	9	AA423158
	40	81.6	13.0	631	12	BG080418
	41	80.2	12.8	756	13	BI278559
	42	78	12.4	715	17	AQ506619
	43	77.2	12.3	124	9	A1182435
	44	77.2	12.3	648	12	BF393825
	45	77	12.3	513	9	AA418441

ALIGNMENTS

RESULT 1  
BF603486  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS

BF603486  
269014 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.  
BF603486  
BF603486.1 GI:11701143  
EST.  
cow.  
Bos taurus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
Bovidae; Bovinae; Bos.  
1 (bases 1 to 504)  
Smith, P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,  
Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett  
, G.L., Heaton, M.P., Laegreid, W., Rohrer, G.A., Chitko-McKown, C.G.,  
Perte, G., Holt, I., Karameycheva, S., Liang, F., Quackenbush, J. and  
Keele, J.W.

TITLE  
JOURNAL  
MEDLINE  
COMMENT

Sequence evaluation of four pooled-tissue normalized bovine cDNA  
libraries and construction of a gene index for cattle  
Genome Res. 11 (4), 626-630 (2001)  
21180013  
Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: smith@email.marc.usda.gov  
Single pass sequencing. Bases called and alt\_trimmed with phred

v0.980904.e. Vector identified by cross\_match with the -minscore 18 and -minmatch 12 options.

PCR Primers  
FORWARD: AGGAACAGCTATGACCAT  
BACKWARD: GTTTTCCCAGTCACGACG  
Plate: 48 row: N column: 13  
Seq primer: ATTTAGTGCACATATAG.

FEATURES

Location/Qualifiers  
1. .504

/organism="Bos taurus"  
/db\_xref="taxon:9913"  
/clone\_lib="MARC 3BOV"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/note="Vector: pCMV SPORT6; Site\_1: NotI; Site\_2: SalI;  
Library made from pooled tissue from marrow, alveolar  
macrophage, ovary, fetal semitendinosus muscle, and fetal  
longissimus muscle."

BASE COUNT 148 a 92 c 126 g 137 t 1 others

Query Match 53.0%; Score 332.4; DB 12; Length 504;  
Best Local Similarity 83.2%; Pred. No. 2.1e-92;  
Matches 392; Conservative 0; Mismatches 72; Indels 7; Gaps 1;

Qy 164 CTGAACCTACACTATATCATTCAGTCTCACCTGCTTCAGTGAAGGACAAAGGTGCCAG 223

Db 8 CTGAACCTGCACACACACATTCAAATCAACCTGCTTCAGTGAAGGACAAAGGTGCACAN 67

Qy 224 CC-----TGGGATGTTGCCAGCTCTTGGGAAGTCATTTGGTTCACGTTCTACTTC 276

Db 68 GAAAGATTTGGGATGTTGCCAGGTACCTTGGGAAGCCGTTTGGTCCAGCTGCTACTTT 127

Qy 277 ATTTCCAGTGAAGAAGGTTTGGTCTAAGAGTGAGCAGAACTGTGTGAGATGGGAGCA 336

Db 128 ATTTCTCTGAAGAGAATTTCTGGGCTAAGAGTGAGCAGAACTGTCATGGGATGGGAGCT 187

Qy 337 CATTTGGTTGTTCACACAGAAGCAGACAGAAATTTCAATTGTCCAGCAGCTGAATGAG 396

Db 188 CACTTGGTGGTGATCAACACAGAAACAGACAGGATTTCAATTATCCAGCAGCTGAATAA 247

Qy 397 TCATTTCTTATTTCTGGGCTTTCAGACCCCAAGGTAATATATATTTGCAATGGATT 456

Db 248 ACATTTCTTATTTCTGGGACTCTCAGACCCCAAGGTAATATATATTTGCAATGGATT 307

Qy 457 GATAAGCACCTTATGAGAAAATGTCAGATTTTGGCACCTAGGTGAGCCCAATCAATCT 516

Db 308 GATCAGACACCTTACAAGGAAATGTCAGATTTTGGCACCAAAATGAACCAACTTTTCT 367

Qy 517 GCAGAGCAATGCTTCAATAGTCTTCTGGAAACCTACAGATGGGGCTGGAATGATTT 576

Db 368 GCAGAGGAATGCTTCAATAGTCTTCTGGGATGGGAGAGGATGGGCTGGAATGATTT 427

Qy 577 ATCTCTGAACCTAGAGGAATTCATATGTCAGATGAATAGATTTTACCTA 627

Db 428 TTCTGTGATCTAAAGGAAATCAATATGTGAGATGAAGAGATTTACCTA 478

RESULT 2  
BF774158 470 bp mRNA linear EST 25-APR-2001

LOCUS 283735 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.

DEFINITION BF774158

ACCESSION BF774158

VERSION BF774158.1 GI:12122058

KEYWORDS EST.

SOURCE COW.

ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovidae; Bovinae; Bos.

1 (bases 1 to 470)

Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,

Casas, E., Wray, J.E., White, J., Cho, J., Fahnenkrug, S.C., Bennett

TITLE

JOURNAL

MEDLINE

COMMENT

,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,  
Perteu,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and  
Keele,J.W.  
Sequence evaluation of four pooled-tissue normalized bovine cDNA  
libraries and construction of a gene index for cattle  
Genome Res. 11 (4), 626-630 (2001)

21180013

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt.trimmed with phred

v0.980904.e. Vector identified by cross\_match with the -minscore 18

and -minmatch 12 options.

PCR Primers

FORWARD: AGGAACAGCTATGACCAT

BACKWARD: GTTTTCCCAGTCACGACG

Plate: 80 row: M column: 16

Seq primer: ATTTAGTGCACATATAG.

Location/Qualifiers

1. .470

/organism="Bos taurus"

/db\_xref="taxon:9913"

/clone\_lib="MARC 3BOV"

/tissue\_type="pooled"

/lab\_host="DH10B"

/note="Vector: pCMV SPORT6; Site\_1: NotI; Site\_2: SalI;

Library made from pooled tissue from marrow, alveolar

macrophage, ovary, fetal semitendinosus muscle, and fetal

longissimus muscle."

BASE COUNT 138 a 84 c 122 g 124 t 2 others

ORIGIN

Query Match 51.0%; Score 320; DB 12; Length 470;

Best Local Similarity 82.9%; Pred. No. 1.5e-88;

Matches 379; Conservative 0; Mismatches 71; Indels 7; Gaps 1;

Qy 165 TGAACCTACACTCATATCATTCAGTCTCACCTGCTTCAGTGAAGGACAAAGGTGCCAGC 224

Db 13 TGAACCTGCACACACACCATTCAAATCAACCTGCTTCAGTGAAGGACAAAGGTGCACANG 72

Qy 225 C-----TGGGATGTTGCCAGCTCTTGGAGTCAATTTGGTTCAGTCTACTTCA 277

Db 73 AAAAGATTTGGGATGTTGCCAGGTACCTGGAAGCCGTTTGGTTCAGCTCTACTTTA 132

Qy 278 TTTCCAGTGAAGAGAGGTTTGGTCTTAAGAGTGAGCAGAACTGTGTGAGATGGGAGCAC 337

Db 133 TTTCTCTGAAGAGAAATTTCTGGGCTAAGAGTGAGCAGAACTGCATTTGGATGGGAGCTC 192

Qy 338 ATTTGGTGTGTTCACACAGAAGCAGACGAAATTTCAATTCCTCCAGCAGCTGAATGAGT 397

Db 193 ACTTGGTGGTGATCAACACAGAAACAGAGCAGGATTTCAATTCAGCAGCTGGAATAAAA 252

Qy 398 CATTTTCTTATTTCTGGGGCTTTCAGACCCACAGGTAATAATAATTTGGCAATGGATTG 457

Db 253 CATTTTCTTATTTCTGGGACTCTCAGACCCACAGGGAATGGCAATGGATTG 312

Qy 458 ATAAGACACCTTATGAGAAAATATGTCAGATTTTGGCACCTAGGTGAGCCCAATCTCTG 517

Db 313 ATCAGACACCTTACAGGAAATATGTCAGATTTTGGCACCACCAAAATCAACCAACTTTCTG 372

Qy 518 CAGAGCAATGTGCTTCAATAGTCTTCTGGAAACCTTACAGGATGGGGCTGGGAATGATGTTA 577

Db 373 CAGAGGAATGTGCTTCAGTGTGTTTCTGGATGGGAGAGGATGGGCTGGAATGATGTTT 432

Qy 578 TCTGTCAAACTAGAAGAAATTCAAATATGTGAGATGAA 614

Db 433 TCTGTGATTTCTAAAGGAAATCAATATGTGAGATGAA 469

RESULT 3

BI696642  
 LOCUS 603348532F1 NCI\_CGAP\_Mam2 Mus musculus cdna clone IMAGE:5376247 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BI696642  
 VERSION BI696642  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NIH-MGC http://mgc.nci.nih.gov/  
 1 (bases 1 to 971)  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-r@mail.nih.gov  
 Tissue Procurement: Gilbert Smith, Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L14M1957 row: p column: 08  
 High quality sequence start: 5  
 High quality sequence stop: 777.  
 FEATURES  
 source  
 1..971  
 /organism="Mus musculus"  
 /strain="FVB/N-3"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5376247"  
 /clone.lib="NCI\_CGAP\_Mam2"  
 /tissue\_type="tumor, biopsy sample"  
 /dev\_stage="5 months"  
 /lab\_host="DH10B"  
 /note="organ: mammary; Vector: pCMV-SPORT6; Site\_1: SalI;  
 Site\_2: NotI; Cloned unidirectionally. primer: Oligo dt.  
 Library constructed by Life Technologies. Investigator  
 providing samples: Gilbert Smith, NIH"  
 BASE COUNT 258 a 222 c 229 g 262 t  
 ORIGIN  
 Query Match 49.4%; Score 310; DB 13; Length 971;  
 Best Local Similarity 73.8%; Pred. No. 2.7e-85;  
 Matches 450; Conservative 0; Mismatches 150; Indels 10; Gaps 4;  
 QY 10 GAGCAGAACCTCAAGTACAGAGAAAGAGCGTCGGTTGCCTCAGACTCTGGTCTG 69  
 Db 149 GTGCAGGAAGACAAATCCAGGGAAGGAGTCTGCTGGACCTCAGACTCTGGTCACT 208  
 QY 70 GCTGGGATTTCCATTGSCACTCCTCAGTCTGCTTCATTGTGAGCTGTGTAGTAACCTAC 129  
 Db 209 GCTGTGATTTCCATGCTTACTCTTGAGTACCTCTTTTCATTGCGAGCTGTGTGCTGACTTAC 268  
 QY 130 CATTTTACATATGTTGAACCTGGCAAAAGGCTGTCTGAACCTACATCATATCAATCAACT 189  
 Db 269 CAATTTATATGAGCACGCCAGTGAAGACTATATGAACCTTCACATACATCACTCCAGT 328  
 QY 190 CTCACCTGCTTCAGTGAAGGACAAAGGTGCCAG-----CCTGGGGATGTTGCCACAGT 243  
 Db 329 CTCACCTGCTTCAGTGAAGGACTATGGTGTGAGAAAATCTGGGATGCTGCCAAT 388  
 QY 244 TCTTGGAGATCATTTGGTTCAGTTGCTACTTCATTTCAGTGAAGAGAGGTTTGGTCT 303  
 Db 389 CACTGGAAGTCAATTTGGTCCAGCTGCTACCTCATTTCTACCAAGGAGAACTCTTGAGC 448  
 QY 304 AAGAGTGACCAACTGTGTCGAG-ATGGGACCAATTTGGTGTGTTCTCAACACAGAGC 362  
 Db 449 ACCAGTGAGCAAGACTGTGTCAGATGGGGCTCATCTGGTGGTGAATCAATCACTGAAGC 508  
 QY 363 AGAGCAGAAATTTTCATTGTGCCAGCAGCTGAATGAGTCAATTTCTTAATTTTCTGGGCTTTC 422

[illegible]





201203, P. R. China  
Tel: 86-21-50801919(ex.45)  
Fax: 86-21-50801922  
Email: hanzg@chgc.sh.cn  
This clone is available at CHCC in Shanghai.

FEATURES  
source

Location/Qualifiers  
1. .582

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="DCBBA12"  
/clone.lib="DCB"  
/cell\_type="dendritic cells"  
/dev\_stage="mature"  
/lab\_host="BM25.8"  
/note="Vector: pTriplex2; Site\_1: sf1A; Site\_2: sf11B"  
171 t

BASE COUNT 162 a 110 c 139 g 171 t  
ORIGIN

Query Match 26.5%; Score 166; DB 10; Length 582;  
Best Local Similarity 67.8%; Pred. No. 1.8e-40;  
Matches 232; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 215 AGGTGCGACGCTGGGATCTGCCAGCTCTTGGAGTCAATTTGGTCCAGTTGCTACT 274  
Db 4 AAGAGACACCTGGAGCGTGCCAAAGATTTGAAGTCATTTAGTTCACACTGCTACT 63  
QY 275 TCATTTCCAGTGAAGAAGGTTTGGTCTAAGAGTGAGCAGAACTGTGTTGAGATGGGAG 334  
Db 64 TTATTTCTACTGAATCAGCATCTTGGCAAGACAGTGAGAAGGACTGTGCTAGAATGGAGG 123  
QY 335 CACATTTGGTTGTTCAACACAGAGCAGCACAATTTCAATTCAGCAGCTGAATG 394  
Db 124 CTCACCTGCTGTGTAAACACTCAAGAAGACAGCAGGATTTCACTTCAGAACTCTGCAGG 183  
QY 395 AGTCATTTTCTATTCTTCGGGGCTTTCAGACCCACAAGGTAATAATTTGGCAATGGA 454  
Db 184 AAGAATCTGCTATTCTTGGGGCTCTCAGATCCAGAGGTCAGGCACATTTGGCAATGGG 243  
QY 455 TTGATGAACACCTATGAGAAATATGAGATTTTGGCACTAGTGTAGGCCCAATCAAT 514  
Db 244 TTGATCAGACACATACATGAATGAAGTTCACATTTCCGATCTGCGATCCAGTCCAGCCAGTGATC 303  
QY 515 CTCGACAGCAATGCTTCAATAGTCTCTGGAACCTACAG 556  
Db 304 CCAATGAGCGCTGCGTGTGCTGCTAAATTTTCGTAATATCATCCG 345

RESULT 9  
BF152383

LOCUS BF152383 514 bp mRNA linear EST 29-DEC-2000  
DEFINITION uz32f08.y1 NCI\_CGAP\_Man5 Mus musculus cdna clone IMAGE:3670791 5', similar to TR:Q9Q215 Q9Q215 DENDRITIC CELL IMMUNORECEPTOR. ;, mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE

BF152383  
BF152383.1 GI:11033778

EST.  
house mouse.  
Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 514)

REFERENCE

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL

COMMENT Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaps-remail.nih.gov  
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys  
cdna Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
Cloned by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

image.llnl.gov/image/html/iresources.shtml

MGI:1431559  
Seq primer: -40RP from Gibco  
High quality sequence stop: 436.

FEATURES  
source

Location/Qualifiers  
1. .514

/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:3670791"  
/clone.lib="NCI\_CGAP\_Man5"  
/tissue\_type="tumor, gross tissue"  
/dev\_stage="7 months"  
/lab\_host="DH10B"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site\_1: SalI; Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Robin Humphreys, NIH"

BASE COUNT 161 a 96 c 117 g 140 t  
ORIGIN

Query Match 24.1%; Score 151.4; DB 12; Length 514;  
Best Local Similarity 64.1%; Pred. No. 6.3e-36;  
Matches 262; Conservative 0; Mismatches 141; Indels 6; Gaps 2;

QY 222 AGCTGTGGGATGTGGCCAGCTCTTGGAAATCATTTGGTTCCAGTTGCTACTTCAATTC 281  
Db 109 AGTCTGGAGCTGTGGCCAAAGGATTGGAAGCCGTTGGTTCCCTACTGCTACTTCACTTC 168  
QY 282 CAGTGAAGAAGGT---TTGGTCTAAGAGTGAGCAGAACTGTGTGAGATGGAGCACA 338  
Db 169 AACTGATCTGGTGATCTTGGAAATGAGATGAAGAGAACTGCTCCACATGGGTGCTCA 228  
QY 339 TTTGGTTGTGTTCAACACAGCAGCAGCAATTTCAATGTCCAGCAGCTGAATGATC 398  
Db 229 TCTGGTGGTGATCCACAGCCAGGAAGAACAGGATTTCACTACCTGGGATCCTGGACACTGG 288  
QY 399 ATTTTCTATTCTTGGGGCTTTCAGACCCACAAGTAAATAATTTGCAATGGATTGA 458  
Db 289 TACTGCTATTATTATAGGACTTCAATCC--AGGTGATCAACATGGCAATGGATTGA 345  
QY 459 TAAGACACCTTATGAGAAAAATGTGAGATTTTGGCACTAGTGTAGGCCCAATCAATCTTCG 518  
Db 346 TCAGACACCGTAGATGATAACACCATCTCTGGCAAAAGTGAGCTAGCAGTAGACAA 405  
QY 519 AGACAATGTGCTCAATAGTCTTCTGGAACCTACAGGATGGGCTGGAATGATGTAT 578  
Db 406 TGAACAGTGTGTTATATAATAATCATGCTCAGAGTACTGGATGGGCTGGAGTGATATCC 465  
QY 579 CTGTGAACTAGAAGGAATTCATATGTGAGATGAATAAGATTTTACTTA 627  
Db 466 TTGCAGTGATAAACAACAACATCAATTTGCCATGTGAAAAAATATACTTA 514

RESULT 10  
AI391056

LOCUS AI391056 558 bp mRNA linear EST 15-MAR-2000  
DEFINITION mc13h05.y1 Soares mouse p3NMF19.5 Mus musculus cdna clone IMAGE:348441 5', similar to SW:MANR\_HUMAN P22897 MACROPHAGE MANNNOSE RECEPTOR PRECURSOR. ;, mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE

AI391056  
AI391056.1 GI:4217063

EST.  
house mouse.  
Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 558)  
REFERENCE Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter

.E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,  
 Waterston, R. and Wilson, R.  
 The WashU-NCI Mouse EST Project 1999  
 Unpublished (1999)  
 Contact: Marra M/WashU-NCI Mouse EST Project 1999  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: mouseest@watson.wustl.edu  
 This clone is available royalty-free through LLNL ; contact the  
 IMAGE Consortium (infoimage.llnl.gov) for further information.  
 This read is a RESEQUENCE of a previously sequenced mouse clone  
 This read has been verified (found to hit its original self in the  
 correct orientation)  
 MGI:220241  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 491  
 POLYA=No.

FEATURES	source	BASE COUNT	ORIGIN
Location/Qualifiers			
1. .558			
/organism="Mus musculus"			
/db_xref="taxon:10090"			
/clone="IMAGE:348441"			
/clone_lib="Soares mouse p3NWF19.5"			
/dev_stage="19.5 dpc total fetus"			
/lab_host="DI10B (ampicillin resistant)"			
/note="Vector: pT73D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCACTCTGAAGTCGGACGCCGCAATTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Fátima Bonaldu. RNA was kindly provided by Dr. Minoru Ko (Wayne State University)."			
171 a	108 c	129 g	150 t

Query Match	24.1%	Score 151.4	DB 9	Length 558	
Best Local Similarity	64.1%	Pred. No. 6.5e-36			
Matches 262	Conservative	0	Mismatches 141	Indels 6	Gaps 2
222	QY	AGCCTGGGAGTTGGCCACGCTTCTGGAAGTCATTTGGTTCAGTGTGCTACTTCATTTC	281		
76	Db	AGTCTGGAGCTTTGGCCCAAGGATTGGAAGCGTTTGGTCTCTACTGCTACTTCACATTC	135		
282	QY	CAGTGAAGAGAAGT- - - TTGGTCTAAGAGTGAGCAGAACTGTGTGGATGGGAGCAC	338		
136	Db	AACTGACTTGTGTGCATCTTGGAAATGAGATGAAGGAGAACTGCTCCACATGGGTGCTCA	195		
339	QY	TTTGGTGTGTTCACACAGACAGCAGAGAAATTTTCATTGTCCAGCAGCTGAATGAGTC	398		
196	Db	TCTGTGTGTATCCACAGCCAGGAAGACAGGATTTTCATCAGTGGGATCTCTGGACACTGG	255		
399	QY	ATTTCTTATTTCTTGGGGCTTTTCAGACCCACAAAGATAATAATATGGCAATGGATTGA	458		
256	Db	TACTGCTTATTTTATAGGACTTTTCAAATCC- - - AGTGATCAACAATGGCAATGGATTGA	312		
459	QY	TAGACACCTTATCAGAAAAATCTCAGATTTTGGCACCTTAGTGAAGCCCAATCATCTGTC	518		
313	Db	TCGACACCGTACGATGATTAATACCAATCTCTGGCAAAAGGTGAGCCTAGCAGCTGACAA	372		
519	QY	AGACCAATGTGCTTCAATAGTCTCTCTGGAAACCTACAGGATGGGCTGGAATGATGTTAT	578		
373	Db	TGAACAGTGTGTTATAATAAATCATCGTCAGAGTACTGGATGGGCTGGAGTGATATCCC	432		
579	QY	CTGTGAAACTAGAGGAATTCATATGTGAGATGAATAGATTTTACCTA	627		
433	Db	TTGCAGTGTAAACAGAACTCAATTTGGCATGTGAAAAAATATATCTTA	481		

RESULT 11	
BE650539	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
MEDLINE	
COMMENT	

FEATURES SOURCE

BE650539 561 bp mRNA linear EST 06-SEP-2000  
 UI-M-BH2.3-aqf-f-04-0-UI.r1 NIH\_BMAP\_M.S3.3 Mus musculus cDNA clone  
 UI-M-BH2.3-aqf-f-04-0-UI 5', mRNA sequence.  
 BE650539  
 BE650539.1 GI:9976363  
 EST.  
 house mouse.  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 561)  
 Bonaldo, M.F., Lennon, G. and Soares, M.B.  
 Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 Genome Res. 6 (9), 791-806 (1996)

2704147  
Contact: Chin, H  
National Institute of Mental Health  
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
20892-9643, USA  
Tel: 301 443 1706  
Fax: 301 443 9890  
Email: mEST@mail.nih.gov  
CDNA Library Preparation: M.B. Soares Lab Clone distribution:  
Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It  
should be noted that Bento Soares is generating a small number of  
additional specialized non-redundant arrays of BMAP cDNAs whose  
availability will be considered under appropriate and limited  
collaborative arrangements  
Seq primer: M13 Reverse.

```

1. 561
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-BH2.3-af-f-04-0-UI"
/clone.lib="NIH_BMAP_M_S3.3"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pTT30-Pac (Pharmacia) with a modified
polylinker. Site.1: Not I; Site.2: Eco RI; The
NIH_BMAP_M_S3.3 library is a subtracted library of a
series, ultimately derived from a mixture of individually
tagged normalized libraries from ten regions of the mouse
brain (cerebellum, brain stems, olfactory bulbs,
hypothalamus, cortex, amygdala, basal ganglia, pineal
gland, striatum, hippocampus) after a series of
subtractions to reduce the representation of cDNAs from
which ESTs had already been generated. The following
serially subtracted libraries were generated in this
process: NIH_BMAP_M_S3.3, NIH_BMAP_M_S2, NIH_BMAP_M_S1.
The subtracted library (NIH_BMAP_M_S3.3) was constructed
as follows: PCR-amplified cDNA inserts from NIH_BMAP_M_S2
clones from which 3' ESTs had been derived was used as a
driver in a hybridization with the NIH_BMAP_M_S2 library
in the form of single-stranded circles. The remaining
single-stranded circles (subtracted library) was purified
by hydroxapatite column chromatography, converted to
double-stranded circles and electroporated into DH10B
bacteria (Life Technologies) to generate the
NIH_BMAP_M_S3.3 library. This procedure has been
previously described (Bonaldo, Lennon and Soares, Genome
Research 6:791-806, 1996)"
155 a 117 c 144 g 145 t

```

BASE COUNT  
ORIGIN

Query Match 22.6%; Score 141.8; DB 10; Length 561;  
Best Local Similarity 60.9%; Pred. No. 6.5e-33;  
Matches 249; Conservative 0; Mismatches 157; Indels 3; Gaps 1;



/note="data source:SPTR, source key:Q9QZ15, evidence:ISS  
putative  
similar to DENDRITIC CELL IMMUNORECEPTOR"  
/codon\_start=1  
/protein\_id="BAB25260.1"  
/db\_xref="GI:12841568"  
/db\_xref="MGD:MGI:1917060"  
/translation="MYOERQLOQKAVSWLSRLWSAAVISILLSTCTFIASCVDKVMWSC  
CPDKWLFGSHCVLPVTFVSSAWNKSEENRGMGAHLVVIHQREODPTIGLIDIIHA  
AFVIGLWDTHCHBQWQVDTPEEVSVPFHNGEPSSDNEKCVTVYRRNIGWGNDIS  
CNLKQSVCKMKNL"  
BASE COUNT 174 a 129 c 157 g 171 t  
ORIGIN

Query Match 22.3% Score 140: DB 11: Length 631;  
Best Local Similarity 60.6% Pred. No. 2.5e-32;  
Matches 268: Conservative 0; Mismatches 165; Indels 9; Gaps 2;  
QY 192 CACCTGTTTCAGTGAAGGACAAAGTCCAGGCTGGGGATGTTGCCAGCTTCTTGGAA 251  
||||| ||||| || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 183 CACCTGTTTCATTCGAGCTGTGTGGACAAAGTCTGGAGCTGTGCCAAAGGATTTGGAA 242  
QY 252 GTCATTTGGTTCAGTTCCTACTTCATCTCCAGTGA-----AGACGAAGTTTGGTCTAA 305  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 243 GCTATTTGGTTCACACTGCTACTTGGTTCACAGTCTTTTTCATCAGCATCTTTGGAA 302  
QY 306 GAGTGACGACAACTGTCTTGAGATGGGACGACATTTGGTTGCTTCAACACAGACGAGA 365  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 303 GAGTGACGACAACTGCTCCGATGGTGTCTATCTAGTGGTGATCCATAGCCAGGAGGA 362  
QY 366 GCAGAATTTCAATGTCAGGAGTGAATGAGTCAATTTCTTATTTTCTGGGGCTTTTTCAGA 425  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 363 GCAGGATTTCACTACTGGGATCTTGGACATTCATCTGCTGCTTATTTATAGGTTGTGGGA 422  
QY 426 CCCACAGGTAATAATTTGGCAATGGATTGATAGACACCTTATGAGAAAATGTCAG 485  
Db 423 TAC---AGGCGATCGGCAATGGCAATGGTGTGATGATGATGATGATGATGATGATGATG 479  
QY 486 ATTTTGGCAGCTAGTGCAGCCCAATCATCTTCGACGACATGTCCTCAATAGTCTCTG 545  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 480 ATCTGCGACAAATGGTGCAGCCAGCAGTGCACATGAAATGTTGTACAGTATATTACCG 539  
QY 546 GAAACATACAGGATGGGGCTGGAAATGATGTTATCTGTGAACATAGAGGAATTCATATG 605  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 540 TCGGAATATTGGATGGGCTGGAATGATATCTCTTGCAATCTTAAACAGAGTCAGTTTG 599  
QY 606 TGAGATCAATAAGATTTACCTA 627  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 600 TCAGATGAAGAAAATAAACTTA 621

RESULT 13  
AK014135  
LOCUS AK014135 1200 bp mRNA linear HTC 19-JAN-2002  
DEFINITION Mus musculus 13 days embryo head cDNA, RIKEN full-length enriched library, clone:3110037K17:similar to DENDRITIC CELL IMMUNORECEPTOR, full insert sequence.  
ACCESSION AK014135  
VERSION AK014135.1 GI:12851811  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1  
AUTHORS Carninci,P. and Hayashizaki,Y.  
TITLE High-efficiency full-length cDNA cloning  
JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
MEDLINE 99279253  
PUBMED 10349636  
REFERENCE 2

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
11042159  
3  
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuana,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaki,S., Inoue,K., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861  
4  
Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I., Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R., Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T., Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H., Kuehl,P., Lewis,S., Matsuo,Y., Nikaide,I., Pesole,G., Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tonita,M., Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N., Carindrelli,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C., Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D., Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seva,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S. and Hayashizaki,Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)  
21085660  
11217851  
5 (bases 1 to 1200)  
Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F., Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J., Saito,H., Saito,R., Sakai,K., Sakai,K., Sano,H., Sasaki,D., Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Soabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.  
Direct Submission  
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)  
Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Sciences Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to



Search completed: February 17, 2003, 11:47:15  
Job time : 607.362 secs

ACCESSION AW242656  
VERSION AW242656.1 GI:6576410  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 666)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
DNA Sequencing by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
Seq primer: -400P from Gibco  
High quality sequence stop: 422.

FEATURES  
source  
1..666  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2691346"  
/clone\_lib="NCI-CGAP\_Kid11"  
/lab\_host="DH10B"  
/notes="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with  
a modified polylinker; Site\_1: Not I; Site\_2: Eco RI;  
Plasmid DNA from the normalized library NCI-CGAP\_Kid3 was  
prepared, and ss circles were made in vitro. Following HAP  
purification, this DNA was used as tracer in a subtractive  
hybridization reaction. The driver was PCR-amplified cDNAs  
from a pool of 5,000 clones made from the same library  
(clones 1322376-1323911, 1456007-1456775, and  
1500552-1502855). Subtraction by Bento Soares and M.  
Fatima Bonaldo."

BASE COUNT 201 a 147 c 124 g 194 t

Query Match 21.5% Score 135; DB 10; Length 666;  
Best Local Similarity 65.4%; Pred No. 9.2e-31;  
Matches 214; Conservative 0; Mismatches 110; Indels 3; Gaps 1;  
QY 304 AAGAGTGAGCAGAACTGTGTTGAGATGGGAGCACCAATTTGGTTGTGTTCAACACAGAAGCA 363  
Db 663 AACAGTGACAAGGACTGTGCAATAATATGATGCTCCCTTGGTGTGATAAACAACACTCAAGAA 604  
QY 364 GAGCAGAAATTTTCATGTCACGACGCTGAATGAGTCATTTTCTTATTTCTGGGGCTTCA 423  
Db 603 GAGCAGGAAATTTTCATGTTCAAAATCTCAAGAGAATCCGCTTATTTCTGGGGCTTCA 544  
QY 424 GACCCACAAGGTAATAATTTGGCAATGGATTGATAAGACCTTATGAGAAAAATGTC 483  
Db 543 GATCCAGTAGTCAGCGACATTTGGCAATGGTTGATCAGACACCAATACAAATGAAAGTTCC 484  
QY 484 AGATTTTGGCACCTAGGTGAGCCCAATCATCTGCAGAGCAATGCTTCATAGTCTTC 543  
Db 483 ACATTTCTGGCATCCAGCTGAGCTCAGTCAATGATCCCAATGAGCGCTGCTGTGCTAAATTTT 424  
QY 544 TGGAA---ACCTACAGGATGGGCTGGAATGATGTTATCTGTGAAACTAGAGAAGAAATCA 600  
Db 423 CTAATACCCCAAGAGATGGGCTGGAATGATGTTATTTCTTCTGCTCTCAAGGTCA 364  
QY 601 ATATGTGAGATGAATAAGATTTACCTA 627  
Db 363 GTTTGTGAGATGATGAAGATCCACTTA 337





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 17, 2003, 08:46:11 : Search time 1081.37 seconds  
(without alignments)  
16874.395 Million cell updates/sec

Title: US-09-766-511B-52  
Perfect score: 627  
Sequence: 1 atgatgaagagcagcaacc.....agatgaataagatttaccta 627

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:

- 1: gb\_ba:\*
- 2: gb\_htg:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vi:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_mu:\*
- 20: em\_om:\*
- 21: em\_or:\*
- 22: em\_ov:\*
- 23: em\_pat:\*
- 24: em\_ph:\*
- 25: em\_pi:\*
- 26: em\_ro:\*
- 27: em\_sts:\*
- 28: em\_un:\*
- 29: em\_vi:\*
- 30: em\_htg\_hum:\*
- 31: em\_htg\_inv:\*
- 32: em\_htg\_other:\*
- 33: em\_htg\_mus:\*
- 34: em\_htg\_pln:\*
- 35: em\_htg\_rtd:\*
- 36: em\_htg\_mam:\*
- 37: em\_htg\_vrt:\*
- 38: em\_sy:\*
- 39: em\_htgo\_hum:\*
- 40: em\_htgo\_mus:\*
- 41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	373.2	59.5	1227	6	AX155225	AX155225 Sequence
2	373.2	59.5	1227	10	AF240357	AF240357 Mus muscu
3	373.2	59.5	1254	10	BC023008	BC023008 Mus muscu
4	273.4	43.6	851	9	AF325459	AF325459 Homo sapi
5	273.4	43.6	1312	6	AX155223	AX155223 Sequence
6	273.4	43.6	1313	9	AF293615	AF293615 Homo sapi
7	271.8	43.3	1125	10	AF240358	AF240358 Mus muscu
8	269.4	43.0	827	6	AX357481	AX357481 Sequence
9	265.6	42.4	800	6	AX357483	AX357483 Sequence
10	243	38.8	1125	10	AF240359	AF240359 Mus muscu
11	229.8	36.7	739	9	AF325460	AF325460 Homo sapi
12	201.2	32.1	444	6	AX456975	AX456975 Sequence
13	200.4	32.0	402	6	AX456962	AX456962 Sequence
14	192	30.6	1133	9	AF109146	AF109146 Homo sapi
15	192	30.6	1195	6	AX400064	AX400064 Sequence
16	192	30.6	1271	9	HSA133532	AJ133532 Homo sapi
17	191.2	30.5	615	9	AF200738	AF200738 Homo sapi
18	190.4	30.4	1305	9	AF067800	AF067800 Homo sapi
19	181	28.9	1308	9	AF328684	AF328684 Homo sapi
20	176	28.1	582	6	AX456965	AX456965 Sequence
21	172.8	27.6	753	6	AX456967	AX456967 Sequence
22	166.8	26.6	1013	6	AX456966	AX456966 Sequence
23	151.4	24.1	558	6	AX456968	AX456968 Sequence
24	151.4	24.1	1192	10	BC034893	BC034893 Mus muscu
25	148	23.6	117296	9	AC092865	AC092865 Homo sapi
26	144.4	23.0	109320	9	AC092746	AC092746 Homo sapi
27	141.8	22.6	561	6	AX456972	AX456972 Sequence
28	135.6	21.6	1037	10	AF387099	AF387099 Mus muscu
29	135	21.5	448	6	AX456973	AX456973 Sequence
30	135	21.5	666	6	AX456971	AX456971 Sequence
31	134.6	21.5	938	10	BC006623	BC006623 Mus muscu
32	134	21.4	568	6	AX456969	AX456969 Sequence
33	134	21.4	1418	10	MM0133533	AJ133533 Mus muscu
34	130.2	20.8	403	6	AX456970	AX456970 Sequence
35	127	20.3	918	10	AF061272	AF061272 Mus muscu
36	121.8	19.4	1973	9	BC032313	BC032313 Homo sapi
37	120.2	19.2	753	9	AF411850	AF411850 Homo sapi
38	117.4	18.7	936	9	BC000715	BC000715 Homo sapi
39	117.4	18.7	968	6	AX083498	AX083498 Sequence
40	117.4	18.7	997	6	AX375956	AX375956 Sequence
41	117.4	18.7	2143	9	AB024718	AB024718 Homo sapi
42	105.2	16.8	110000	2	AC092450_2	Continuation (3 of
43	96.6	15.4	559	6	AX456974	AX456974 Sequence
44	96.6	15.4	5577	2	AC107915	AC107915 Homo sapi
45	96.6	15.4	110000	2	AC092450_2	Continuation (3 of

ALIGNMENTS

RESULT 1  
AX155225  
LOCUS AX155225  
DEFINITION Sequence 3 from Patent WO0136487.  
ACCESSION AX155225  
VERSION AX155225.1 GI:14536703  
KEYWORDS  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 1227)  
AUTHORS Schmitz,J., Dzionek,A. and Buck,D.W.  
TITLE Antigen-binding fragments specific for dendritic cells,  
compositions and methods of use thereof antigens recognized thereby

JOURNAL and cells obtained thereby  
Patent: WO 0136487-A 3 25-MAY-2001;  
Miltényi Biotec GmbH (DE)  
Location/Qualifiers  
1..1227  
Source  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
146..775  
/note="coding sequence of mouse Dectin-2"  
/codon\_start=1  
/protein\_id="CAC42676.1"  
/db\_xref="GI:14536704"  
/translation="MVORQSGKGVCTLRWSAAVISMLLLSTCFIASCVVYQFI  
MDPSRRIVELHTVHSLTFCFSEGTVMSEKMMGCCPNHWSFGSSCYLISIKENFWST  
SEONCVQMGALHVVINTEAEQNFITQOLNESLSYFLGLSDPOGNGKQWIDDTFESQN  
VRFWHPEPNLPERCVSIYVWNPFSKMGNDVFCDSKHSNICEKMKIYL"  
BASE COUNT 335 a 271 c 276 g 345 t  
ORIGIN  
Query Match 59.5%; Score 373.2; DB 6; Length 1227;  
Best Local Similarity 76.1%; Pred. No. 5.9e-100;  
Matches 475; Conservative 0; Mismatches 143; Indels 6; Gaps 1;  
QY 10 GAGCAGCAACCTCAAGTAGACAGAGAAAGAGCGTGGTTGTCCTCGAGACTCTGGTCTGTG 69  
Db 149 GTCCAGGAAAGACAATCCCAAGGAAGGAGTCTGCTGGACCTGAGACTCTGGTCAAGT 208  
QY 70 GCTGGGATTTCCATTCACACTCCTCAGTGTCTGCTTCATTTGAGCTGTGTAGTAACCTTAC 129  
Db 209 GCTGTGATTTCCATTCAGTGTCTGCTTCATTTGAGCTGTGTAGTAACCTTAC 268  
QY 130 CATTTTACATATGTTGAACTGGCAAAAGGCTGTCTGAACTACACTCATATCATTCACAACT 189  
Db 269 CAATTTATATGACAGCCCAAGGAGTCTGCTGGACCTGAGACTCTGGTCAAGT 328  
QY 190 CTCACCTGCTTCAGTGAAGGACAAAGGTCGCAG-----CCTGGGAGTGTGCCAGCT 243  
Db 329 CTCACCTGCTTCAGTGAAGGAGTGTGGTGCAGAAAAAATGTGGGATGCTGCCCAAT 388  
QY 244 TCTTGGAGTCATTTGGTCCAGTGTCTGCTTCATTTCCAGTGAAGAGAAAGTTGGTCT 303  
Db 389 CACTGGAAGTCAATTTGGTCCAGTGTCTGCTTCATTTCCAGTGAAGAGAACTTCTGGAGC 448  
QY 304 AAGAGTGAGCAGAACTGTGTGAGATGGGACACATTTGGTGTGTGTTCACACAGAGCA 363  
Db 449 ACCAGTGAGCAGAACTGTGTGAGATGGGAGTGTCTGCTGGAGTGTGTGAGTCAAGT 508  
QY 364 GAGCAGAAATTTCAATTCAGCAGCTGAATGAGTCATTTCTTATTTCTGGGCTTCA 423  
Db 509 GAGCAGAAATTTCAATTCAGCAGCTGAATGAGTCATTTCTTATTTCTGGGCTTTCG 568  
QY 424 GACCCACAAGGTAAATAATTTGGCAATGGATTTGATAGTAAACACAACTTATGAGAAAAATGTC 483  
Db 569 GATCCACAAGGTAAATAATTTGGCAATGGATTTGATAGTAAACACAACTTATGAGAAAAATGTC 628  
QY 484 AGATTTTGGCACCTAGTGTGAGCCCAATCAATTCGACAGACAATGTCTTCAATAGTCCTTC 543  
Db 629 AGGTTCTGGCACCCCATGAACCAATCTTCCAGAGAGAGGGGTGTCTTTCAATAGTTTAC 688  
QY 544 TGGAAACCTACAGATGGGGCTGGAATGATGTTATCTGTGGAACACTGAGAGGAATTCATA 603  
Db 689 TGGAAATCTTCAAGATGGGGCTGGAATGATGTTTCTGTGATAGTAAACACAACTTCAATA 748  
QY 604 TGTGAGATGAATAGATTTACCTA 627  
Db 749 TGTGAAATGAAGAAGATTTACCTA 772  
RESULT 2  
AF240357 1227 bp mRNA linear ROD 02-MAY-2000  
LOCUS Mus musculus strain BALB/c dectin-2 alpha isoform mRNA, complete  
DEFINITION cds.

AF240357  
AF240357.1 GI:7677468  
Mus musculus.  
Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
Edelbaum,D., Shen,G.-L., Shikano,S., Ritter,R. III, Zukas,P.,  
Arlizumi,K., Morita,A. and Takashima,A.  
Cloning of a second dendritic cell-associated C-type lectin  
(dectin-2) and its alternatively spliced isoforms  
J. Biol. Chem. (2000) In press  
REFERENCE 2 (bases 1 to 1227)  
AUTHORS Arizumi,K. and Takashima,A.  
TITLE Direct Submission  
JOURNAL Submitted (02-MAR-2000) Dermatology, UT Southwestern Medical  
Center, 5323 Harry Hines Blvd, Dallas, TX 75235, USA  
FEATURES  
Location/Qualifiers  
1..1227  
/organism="Mus musculus"  
/strain="BALB/c"  
/db\_xref="taxon:10090"  
/cell\_line="XS52"  
/cell\_type="dendritic cell"  
146..775  
/note="C-type lectin; type II transmembrane protein;  
dendritic cell-associated expression; C-type lectin motif"  
/codon\_start=1  
/product="dectin-2 alpha isoform"  
/protein\_id="AAE67177.1"  
/db\_xref="GI:7677469"  
/translation="MVORQSGKGVCTLRWSAAVISMLLLSTCFIASCVVYQFI  
MDPSRRIVELHTVHSLTFCFSEGTVMSEKMMGCCPNHWSFGSSCYLISIKENFWST  
SEONCVQMGALHVVINTEAEQNFITQOLNESLSYFLGLSDPOGNGKQWIDDTFESQN  
VRFWHPEPNLPERCVSIYVWNPFSKMGNDVFCDSKHSNICEKMKIYL"  
BASE COUNT 335 a 271 c 276 g 345 t  
ORIGIN  
Query Match 59.5%; Score 373.2; DB 10; Length 1227;  
Best Local Similarity 76.1%; Pred. No. 5.9e-100;  
Matches 475; Conservative 0; Mismatches 143; Indels 6; Gaps 1;  
QY 10 GAGCAGCAACCTCAAGTAGACAGAGAAAGAGCGTGGTTGTCCTCGAGACTCTGGTCTGTG 69  
Db 149 GTCCAGGAAAGACAATCCCAAGGAAGGAGTCTGCTGGACCTGAGACTCTGGTCAAGT 208  
QY 70 GCTGGGATTTCCATTCACACTCCTCAGTGTCTGCTTCATTTGAGCTGTGTAGTAACCTTAC 129  
Db 209 GCTGTGATTTCCATTCAGTGTCTGCTTCATTTGAGCTGTGTAGTAACCTTAC 268  
QY 130 CATTTTACATATGTTGAACTGGCAAAAGGCTGTCTGAACTACACTCATATCATTCACAACT 189  
Db 269 CAATTTATATGACAGCCCAAGGAGTCTGCTGGACCTGAGACTCTGGTCAAGT 328  
QY 190 CTCACCTGCTTCAGTGAAGGACAAAGGTCGCAG-----CCTGGGATGTGCCAGCT 243  
Db 329 CTCACCTGCTTCAGTGAAGGAGTGTGGTGCAGAAAAAATGTGGGATGCTGCCCAAT 388  
QY 244 TCTTGGAGTCATTTGGTCCAGTGTCTGCTTCATTTCCAGTGAAGAGAAAGTTGGTCT 303  
Db 389 CACTGGAAGTCAATTTGGTCCAGTGTCTGCTTCATTTCCAGTGAAGAGAACTTCTGGAGC 448  
QY 304 AAGAGTGAGCAGAACTGTGTGAGATGGGACACATTTGGTGTGTGTTCACACAGAGCA 363  
Db 449 ACCAGTGAGCAGAACTGTGTGAGATGGGAGTGTCTGCTGGAGTGTGTGAGTCAAGT 508  
QY 364 GAGCAGAAATTTCAATTCAGCAGCTGAATGAGTCATTTCTTATTTCTGGGCTTCA 423  
Db 509 GAGCAGAAATTTCAATTCAGCAGCTGAATGAGTCATTTCTTATTTCTGGGCTTTCG 568  
QY 424 GACCCACAAGGTAAATAATTTGGCAATGGATTTGATAGTAAACACAACTTATGAGAAAAATGTC 483  
Db 569 GATCCACAAGGTAAATAATTTGGCAATGGATTTGATAGTAAACACAACTTATGAGAAAAATGTC 628  
QY 484 AGATTTTGGCACCTAGTGTGAGCCCAATCAATTCGACAGACAATGTCTTCAATAGTCCTTC 543  
Db 629 AGGTTCTGGCACCCCATGAACCAATCTTCCAGAGAGAGGGGTGTCTTTCAATAGTTTAC 688  
QY 544 TGGAAACCTACAGATGGGGCTGGAATGATGTTATCTGTGGAACACTGAGAGGAATTCATA 603  
Db 689 TGGAAATCTTCAAGATGGGGCTGGAATGATGTTTCTGTGATAGTAAACACAACTTCAATA 748  
QY 604 TGTGAGATGAATAGATTTACCTA 627  
Db 749 TGTGAAATGAAGAAGATTTACCTA 772  
RESULT 2  
AF240357 1227 bp mRNA linear ROD 02-MAY-2000  
LOCUS Mus musculus strain BALB/c dectin-2 alpha isoform mRNA, complete  
DEFINITION cds.



JOURNAL MEDLINE PUBMED 21426806 11536172

REFERENCE 2 (bases 1 to 851)

AUTHORS Arce, I., Hernanz, P. and Fernandez-Ruiz, E.

TITLE Direct Submission

JOURNAL Submitted (01-DEC-2000) Biologia Molecular, Hospital de la Princesa, Diego de Leon, 62, Madrid 28006, Spain

FEATURES

source

1..851

Location/Qualifiers

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="12"

/map="12p12.3-p13.2"

1..851

/gene="CLECSF11"

27..668

/gene="CLECSF11"

/note="C-type lectin receptor; DLEC; alternatively spliced"

/codon\_start=1

/product="dendritic lectin"

/protein\_id="AAL37358.1"

/db\_xref="GI:17225337"

/translation="MVPDEPPQDRKGLWNFQKLVMSMAVVSILLISVCFVSVVPH NFMYSKTVKRLSKLREYQVHPSLTCVMEGDKIEDWSCPTPTWTSQSCYFTSTGMO SWTKSOKNSVMGADLVVINTREODFIQLNKRNSYFLGLSDPGRRHWMKWDTP YNENVTWHSSEPNLDERCAINFRSEEWGNDHCHVPQKSIKMKKIYI"

BASE COUNT 242 a 169 c 192 g 248 t

ORIGIN

Query Match 43.6%; Score 273.4; DB 9; Length 851;

Best Local Similarity 66.7%; Pred. No. 3.1e-70;

Matches 426; Conservative 0; Mismatches 201; Indels 12; Gaps 2;

QY 1 ATGATGCAAGAGCAGCAACCTCAAAGTACAGAGAAAGAGGCTGTGTCC---TGAGA 57

DB 27 ATGGTGCTTGAAGAAGGCTCAAGACCGAGAGAAAGGACTCTGTGTTCAGTTGAAG 86

QY 58 CTCTGGTCTGGTGGGATTTCCATTGCACCTCCTCAGTCTGTCTTCAATTGTGAGTGT 117

DB 87 GTCTGGTCCATGGCAGTGTATCCATTGCTTCCCTCAGTCTGTCTTCACTGTGAGTTCT 146

QY 118 GTAGTAACATTACATTTTACATATGCTGAAACTGGCAAAAGGCTGTCTGAACTA----- 171

DB 147 GTGGGCTTCACAATTTATGTATAGCAAACTGTCAAGAGCTGTCCAAAGTACGAGAG 206

QY 172 ---CACTCATATCATTCAGTCTACCTGCTTCAAGTGAAGGACAAAGTCCAGCCTGG 228

DB 207 TATCAACAGTATCATCAAGCCTGACCTGGCTCATGGAAGGAAAGACATAGAAGATTGG 266

QY 229 GGATGTGGCCAGCTTCTTGGAGTCAATTTGGTCCAGTGTGCTACTTCAATTCCAGTGAA 288

DB 267 AGCTGTGTGCCAACCCCTTGGACTTTCATTTTCAGTCTAGTGTCTACTTATTTCTACTGGG 326

QY 289 GAGAAGTTTGTCTAAGAGTTCAGCAGAACTGTGTGTAGATGGGAGCACATTTGGTTGTG 348

DB 327 ATGCAATCTTGGACTAAGAGTCAAAAGAACCTGTCTGTATGGGGCTGTATCTGGTGTG 386

QY 349 TTCAACACAGAGCAGAGCAAGTTTCATTTGCCAGCAGCTGAATGAGTCAATTTTCTTAT 408

DB 387 ATCAACACCCAGGGAAGACAGGATTTTCATTCATTCAGAACTCTGAAAGAAATTTCTTCTTAT 446

QY 409 TTTCTGGGGCTTCAGACCCCAAGGTAATATAATTGGCAATGATTGATGAAGACACT 468

DB 447 TTTCTGGGGCTTCAGATCCAGGGGCTGGGACATTTGGCAATGGTGTGACACAGACCA 506

QY 469 TATGAGAAAAATGTCAAGTTTGGCACCTTAGGTGAGGCCCAATTCATCTCCAGAGCAATGT 528

DB 507 TACAATGAANAATGTCAATTTCTGGCACTCAGGTGAACCCAAATACCTTGTATGAGGTTGT 566

QY 529 GCTTCAATAGTCTTCTGGAAACCTACAGGATGGGCTGGAATGATGTATCTGTGAAACT 588

Db 567 GCATAATAAATTCGGTTCTTTCAGAAAGATGGGCTGGAATGACATTGACATTGTCATGTA 626

QY 589 AGAAGGAATTCATATGTGAGATGAATAAGATTACCTA 627

Db 627 CCTCAGAAGTCAATTTGCAAGATGAAGAAGATCTACATA 665

RESULT 5

AX155223

LOCUS AX155223 1312 bp mRNA linear PAT 22-JUN-2001

DEFINITION Sequence 1 from Patent WO0136487.

ACCESSION AX155223

VERSION AX155223.1 GI:14536702

KEYWORDS human.

SOURCE Homo sapiens

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 1312)

AUTHORS Schmitz, J., Dzionek, A. and Buck, D.W.

TITLE Antigen-binding fragments specific for dendritic cells, compositions and methods of use thereof antigens recognized thereby and cells obtained thereby

JOURNAL Patent: WO 0136487-A 1 25-MAY-2001;

FEATURES

Location/Qualifiers

1..1312

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/note="BDCA-2 cDNA sequence"

BASE COUNT 374 a 301 c 297 g 340 t

ORIGIN

Query Match 43.6%; Score 273.4; DB 6; Length 1312;

Best Local Similarity 66.7%; Pred. No. 3.2e-70;

Matches 426; Conservative 0; Mismatches 201; Indels 12; Gaps 2;

QY 1 ATGATGCAAGAGCAGCAACCTCAAAGTACAGAGAAAGAGGCTGTGTGCC---TGAGA 57

DB 491 ATGGTGCTTGAAGAAGGCTCAAGACCGAGAGAAAGGACTCTGTGTTCAGTTGAAG 550

QY 58 CTCTGGTCTGGTGGGATTTCCATTGCACCTCCTCAGTCTGTCTTCAATTGTGAGTGT 117

DB 551 GTCTGGTCCATGGCAGTGTATCCATTGCTCTCAGTCTGTCTTCACTGTGAGTTCT 610

QY 118 GTAGTAACATTACATTTTACATATGCTGAAACTGGCAAAAGGCTGTCTGAACTA----- 171

DB 611 GTGGTGCTTCACAATTTATGTATAGCAAACTGTCAAGAGGCTGTCCAAAGTACGAGAG 670

QY 172 ---CACTCATATCATTCAGTCTACCTGCTTCAAGTGAAGGACAAAGTCCAGCCTGG 228

DB 671 TATCAACAGTATCATCAAGCCTGACCTGGCTCATGGAAGGAAAGACATAGAAGATTGG 730

QY 229 GGATGTGGCCAGCTTCTTGGAGTCAATTTGGTTCAGTGTGCTACTTCAATTCCAGTGAA 288

DB 731 AGCTGTGTGCCAACCCCTTGGACTTTCATTTTCAGTCTAGTGTCTACTTATTTCTACTGGG 790

QY 289 GAGAAGTTTGTCTAAGAGTTCAGCAGAACTGTGTGTAGATGGGAGCACATTTGGTTGTG 348

DB 791 ATGCAATCTTGGAGTCAAGAGTCAAAAGAACCTGTCTGTATGGGGCTGTATCTGGTGTG 850

QY 349 TTCAACACAGAGCAGAGCAAGTTTCATTTGCCAGCAGCTGAATGAGTCAATTTTCTTAT 408

DB 851 ATCAACACCCAGGGAAGAACAGGATTTTCATTCATTCAGAACTCTGAAAGAAATTTCTTCTTAT 910

QY 409 TTTCTGGGGCTTCAGACCCCAAGGTAATATAATTGGCAATGATTGATGAAGACACT 468

DB 911 TTTCTGGGGCTTCAGATCCAGGGGCTGGGACATTTGGCAATGGTGTGACACAGACCA 970

QY 469 TATGAGAAAAATGTCAAGTTTGGCACCTTAGGTGAGGCCCAATTCATCTCCAGAGCAATGT 528

DB 971 TACAATGAANAATGTCAATTTCTGGCACTCAGGTGAACCCAAATACCTTGTATGAGGTTGT 1030

QY	529	GCTTCAATAGTCTTCTGGAACCTACAGATGGGGCTGGAATGATGTTATCTGTGAACT	588
Db	1031	CGGATAATAATTTCCGTTCTTACAGAGATGGGGCTGGAATGACATTCACCTGTCATGTA	1090
QY	589	AGAAGGAATCAATATGTGAGATGAATPAAGATTACCTA	627
Db	1091	CCTCAGAAGTCAATTTGCAAGATGAAGAAGATCTACATA	1129
RESULT 6			
AF293615		1313 bp mRNA linear PRI 14-FEB-2002	
DEFINITION		Homo sapiens blood dendritic cell antigen 2 protein (BDCA2) mRNA, complete cds.	
ACCESSION		AF293615	
KEYWORDS		AF293615.1 GI:17224597	
SOURCE		Homo sapiens.	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS		1 (bases 1 to 1313) Dzionek,A., Sohma,Y., Nagafune,J., Cella,M., Colonna,M., Facchetti,F., Gunther,G., Johnston,I., Lanzavecchia,A., Nagasaka,T., Okada,T., Vermi,W., Winkels,G., Yamamoto,T., Zysk,M., Yamaguchi,Y. and Schmitz,J.	
TITLE		BDCA-2, a novel plasmacytoid dendritic cell-specific type II C-type lectin, mediates antigen capture and is a potent inhibitor of interferon alpha/beta induction	
JOURNAL		J. Exp. Med. 194 (12), 1823-1834 (2001)	
MEDLINE		21615135	
PUBMED		11748283	
REFERENCE		2 (bases 1 to 1313) Sohma,Y., Johnston,I., Dzionek,A., Nagasaka,T., Nagafune,J., Okada,T., Hirano,T., Schmitz,J. and Yamaguchi,Y.	
AUTHORS		Cloning and characterization of blood dendritic cell antigen 2 (BDCA-2), a C-type lectin expressed on plasmacytoid dendritic cells	
TITLE		Unpublished	
JOURNAL		3 (bases 1 to 1313) Sohma,Y., Johnston,I., Dzionek,A., Nagasaka,T., Nagafune,J., Okada,T., Hirano,T., Schmitz,J. and Yamaguchi,Y.	
REFERENCE		Direct Submission	
AUTHORS		Submitted (04-AUG-2000) Miltenyi Biotec GmbH, Friedrich-Ebert-Str. 68, Bergisch Gladbach 51429, Germany	
JOURNAL		Location/Qualifiers	
FEATURES		source	
		1..1313	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
gene		/cell_type="plasmacytoid dendritic cell; DC-2"	
		1..1313	
CDS		/gene="BDCA2"	
		492..1133	
		/gene="BDCA2"	
		/notes="C-type lectin"	
		/codon_start=1	
		/product="blood dendritic cell antigen 2 protein"	
		/protein_id="AAL37036.1"	
		/db_xref="GI:17224598"	
		/translation="MVPEEPQDREKLMWFOLKVMVAWSILLSSVCFVSSVPH NFMYSKTRLSKLEIYQYHPSLQVMEGKDIEDWSCPTPTWFSQSCYPISTGMQ SWTKSGNK-SVMGADLVINTEEQDFTIQNLKRNSSVFLGLSDPGRHRRHMQVDTP YNENVTFHWSGPNLDERCAIINERSSSEWGNIDHCHVPKSKCKMKKIYI"	
BASE COUNT		374 a 302 c 297 g 340 t	
ORIGIN			
Query Match		43.6%; Score 273.4; DB 9; Length 1313;	
Best Local Similarity		66.7%; Pred. No. 3.2e-70;	
Matches		426; Conservative 0; Mismatches 201; Indels 12; Gaps 2;	
QY	1	ATGATGCAAGAGCAGCAACCTCAAGTACAGAGAAAGAGCGGTGTTGCC---TGAGA	57
Db	492	ATGTGCGCTGAAGAAGAGCCTCAAGACCGAGAGAAAGGACTCTGGTGTCCAGTTGAAG	551

QY	58	CTCTGGTGTGGCTGGGATTTCCATTGCACCTCCCTCAGTCGCTTGCATTGCTGAGCTGCT	117
Db	552	GTCGTGTCAGCGCAGTCGTATCCATCTTGTCCCTCAGTGTCTGTTTCACTGCTGAGTTCT	611
QY	118	GTAGTAACCTTACCATTTTACATATGTTGAACTGGCAAAAGGCTGTCTCAACTA-----	171
Db	612	GTGGTGCCTCACAATTTTATGTATACAAAACCTGTCAAGAGGCTGTCCAAGTTACGAGAG	671
QY	172	---CACTCATATCATTCAGTCTCACTCTTCAGTGAAGGAGCAAAAGTCCAGCGCTGG	228
Db	672	TATCAACAGTATCATCCAAGCCTGACCTGCTCATGTAAGAGAACATAGAAGATTGG	731
QY	229	GGATGTGGCCAGCTTCTTGGGAAGTCATTGTTCCAGTTGCTACTTTCATTCCACTGAA	288
Db	732	AGCTGCTGCCCAACCCCTTGGACTTCATTTCAGTCTGATTTGCTACTTTATTTTCTACTGGG	791
QY	289	GAGAAGTTTGGTCTAAGAGTGAGCAGAACTGTGTTGAGATGGGAGCACATTTGGTTGTG	348
Db	792	ATGCAATCTTGGACTAAGAGTCAAAAGAACTGTTCTGTGATGGGGCTCATCTGGTGGTG	851
QY	349	TTCACACAGAAGCAGAGCAGAAATTTCAATTGTCAGCAGCTGAATGAGTCATTTTCTTAT	408
Db	852	ATCAACACCAGGGAGACACAGGATTTCAATTCAGAAATCTGAAAAGAAATTTCTTCTTAT	911
QY	409	TTTCTGGGCTTTCAGACCCACAAAGTAATAATAATTGGCAATGGATTCATGAAGACCT	468
Db	912	TTTCTGGGCTGTGAGATCCAGGGGTGGGACATTTGGCAATGGGTGGACCAGACACCA	971
QY	469	TATGAAAAAATGTCCAGATTTTGGCACCTPAGGTGAGCCCAATCATTTCTGCAGAGCAATGT	528
Db	972	TACAATGAAAATGTCCACATTTCTGGCACTCAGGTGAACCAATTAACCTTGATGAGCGTGT	1031
QY	529	GCTTCAATAGTCTTCTGGAAACCTACAGATGGGCTGGAATGATGTTATCTGGAACCT	588
Db	1032	GGGATAATAAATTTCCGTTCTTTCAGAAAGAAATGGGCTGGAATGACATTCACCTGTCATGTA	1091
QY	589	AGAAGGAATTCATATGTGAGATGAATAAGATTTTACCTA	627
Db	1092	CCTCAGAAGTCAATTTGCAAGATGAAGAAGATCTACATA	1130
RESULT 7			
AF240358		1125 bp mRNA linear ROD 02-MAY-2000	
LOCUS		Mus musculus strain BALB/c dectin-2 beta isoform mRNA, complete cds, alternatively spliced.	
DEFINITION		AF240358	
ACCESSION		AF240358.1 GI:7677471	
VERSION			
KEYWORDS		Mus musculus.	
SOURCE		Mus musculus	
ORGANISM		Mus musculus	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
AUTHORS		1 (bases 1 to 1125) Artizumi,K., Shen,G.-L., Shikano,S., Ritter,R. III, Zukas,P., Edelbaum,D., Morita,A. and Takashima,A.	
TITLE		Cloning of a second dendritic cell-associated C-type lectin (dectin-2) and its alternatively spliced isoforms	
JOURNAL		J. Biol. Chem. (2000) In press	
REFERENCE		2 (bases 1 to 1125)	
AUTHORS		Artizumi,K. and Takashima,A.	
TITLE		Direct Submission	
JOURNAL		Submitted (02-MAR-2000) Dermatology, UT Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235, USA	
FEATURES		Location/Qualifiers	
source		1..1125	
		/organism="Mus musculus"	
		/strain="BALB/c"	
		/db_xref="taxon:10090"	
		/cell_line="XS52"	
		/cell_type="dendritic cell"	
		146..673	
CDS			

/note="C-type lectin; type II transmembrane protein; dendritic cell-associated expression; C-type lectin motif; alternatively spliced form of dectin-2 alpha mRNA"  
/codon\_start=1  
/product="dectin-2 beta isoform"  
/protein\_id="AF67178.1"  
/db\_xref="GI:7677472"

/translation="MVDEROSQGVGVCTLRLLWSAAVISMILLSTCFIASCVBKMMGC  
CPNHWKFGSSCYLSTKENFWSIQNVCVOMGAHLVINTAEQNFITQOLNLSLSY  
FLGLSDPGNKGWQWIDDTFQSNVRWHPNLPPEERCVSIYWNPSKWGNDFEC  
DSKINSICEMKKIYL"

BASE COUNT 305 a 246 c 257 g 317 t

Query Match 43.3%; Score 271.8; DB 10; Length 1125;  
Best Local Similarity 76.6%; Pred. No. 9.4e-70;  
Matches 333; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 193 ACCTGCTTCAGTGAAGGACAAAGGTGCCAGCTGGGGATGTGGCCAGCTTCCTGGGAAG 252  
Db 236 ACCTGCTTCATTCGAGCTGTGTGAAAAAATGTGGGATGTGCCCAATCACTGGAAG 295  
QY 253 TCATTTGGTCCAGTTCCTACTTCATTCCTCAGTGAAGAGGTTTGGTCTAAGAGTCAG 312  
Db 296 TCATTTGGCTCCAGCTTACCTACTTCCTACCAAGGAGAACTTCTGGAGCACCAGTCAG 355  
QY 313 CAGAACTGTGTGAGATGGGACACATTTGGTTGTTCACACAGAAGCAGAGCAGAAAT 372  
Db 356 CAGAACTGTTCAGATGGGGCTCATCTGGTGTGATCAATACTGAAGCGGAGCAGAAAT 415  
QY 373 TTCATTGTCACGAGCTGAATGAGTCATTTCTTTATTTCTGGGGCTTTCTAGACCCACAA 432  
Db 416 TTCATCACCCAGCAGCTGAATGAGTCACCTTTCTTACTTCCCTGGGTCTTTCCGGATCCACAA 475  
QY 433 GCTAATAAATTTGCAATGGATTGATAGACACCTTATGAAAAATGTCAGATTTGG 492  
Db 476 GGTATGGCAAAATGGCAATGGATGGATGATATCTCTTCAGTCAAAATGTCAGGTTCGG 535  
QY 493 CACCTAGGTGAGCCCAATCATTTCTGCAGCAAGCAATGCTTCAATAGTCTCTGGAAGCT 552  
Db 536 CACCCCATGAAACCAATCTCCAGAGAGCGGTGTGTTCAAATAGTTTACTGGAATCCT 595  
QY 553 ACAGATGGGGTGGAAATGATGATTTATCTGTGAACCTAGAAGCAATCAATATGTGAGATG 612  
Db 596 TCGAAATGGGGTGGAAATGATGTTTCTGTGTAGTAAACCAATCAATATGTGAAATG 655  
QY 613 AATAAGATTTACCTA 627  
Db 656 AAGAAGATTTACCTA 670

RESULT 8  
AX357481  
LOCUS AX357481 827 bp DNA linear PAT 13-FEB-2002  
DEFINITION Sequence 19 from Patent WO0172773.  
ACCESSION AX357481  
VERSION AX357481.1 GI:18674538  
KEYWORDS .  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1  
AUTHORS Werner,G., Phares,W., Jaritz,M., Lapp,H. and Kalthoff,F.S.  
TITLE Organic compounds  
JOURNAL Patent: WO 0172773-A 19 04-OCT-2001;  
Novartis AG (CH)

FEATURES  
source Location/Qualifiers  
1..827  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 242 a 164 c 188 g 233 t

ORIGIN

Query Match 43.0%; Score 269.4; DB 6; Length 827;  
Best Local Similarity 66.8%; Pred. No. 4.8e-69;  
Matches 419; Conservative 0; Mismatches 196; Indels 12; Gaps 2;

QY 13 CAGCAACCTCAAGTACAGAGAAAGAGCTGGTCTGCC---TGAGACTCTGGTCTGG 69  
Db 28 CAGAGGCTTCAAGACCGAGAAAGACTCTGGTGGTTCAGTTGAAGGTCTGGTCCATG 87  
QY 70 GCTGGGATTTCCATTGCACCTCCCTCAGTGTGCTTCATTGTGAGCTGTGTAGTAACATAC 129  
Db 88 GCAGTCGTATCCATCTTGTCTCCTCAGTGTCTGTTTCACTGTGAGTCTGTGGTGCCTCAC 147  
QY 130 CATTTACATATGTTGAACTGGCAAAAGCTGCTCTGAACCTA-----CACTCATAT 180  
Db 148 AATTTTATGATAGCAAAAGCTGTCAAGAGGCTGTCCAAGTTACGAGAGTATCAACAGTAT 207  
QY 181 CATTCAGTCTCACCCTCAGTGAAGGACAAAGTGCCACCTTGGGATGTTGCCCA 240  
Db 208 CATCCAAGCCTGACCTCGCTCATGGAAGGAAAGCATAGAAGATTGGAGCTGCTGCCCA 267  
QY 241 GCTTCTTGGAAAGTCATTGGTTCCTCAGTGTCTACTTTCATTTCCTCAGTGAAGAGGTTTGG 300  
Db 268 ACCCTTGGACTTCATTTCAGTCTAGTGTCTACTTTATTCTACTGGGATGCAATCTGG 327  
QY 301 TCTAAGAGTGAAGCAACTGTGTGAGATGGGACACATTTGGTGTGTTCAACACAGAA 360  
Db 328 ACTAAGAGTCAAAAGAACTGTCTGTGTGGGGCTGATCTGGTGTGATCAACACAGG 387  
QY 361 GCAGAGCAGAAATTCATTGCTCCAGCAGCTGAATGAGTCATTTCTTCTTATTTCTGGGGCTT 420  
Db 388 GAAGAACAGGATTCATTCATTCAGAATCTGAAAAGAAATTCCTTTATTTCTGGGGCTG 447  
QY 421 TCAGACCCACAAGTAAATAATTTGGCAATGGATTGATAAGACACCTTATGAGAAAAAT 480  
Db 448 TCAGATCCAGGGGTGGCCACATTTGGCAATGGTGTGACCAGACACCATACATGAAT 507  
QY 481 GTCAGATTTTGGCACCTTAGTGGAGCCCAATCATTTCTGACAGCAATGTGTTCAATAGTC 540  
Db 508 GTCACATTTCTGGCAGCTCAGTGAACCAATACCTTGATGAGCGGTGTGCGGATAATAAT 557  
QY 541 TTCTGGAAACCTACAGATGGGGCTGGAATGATGTTATCTGTGAAACTAGAAGGAATTC 600  
Db 568 TTCCTGTTTTCAGAAAGATGGGGCTGGAATGACATGACATGCTATGCTACCTCAAGAGTCA 627  
QY 601 ATATGTGAGATGAATAAGATTTACCTA 627  
Db 628 ATTTGCAAGATGAAGAAGATCTACATA 654

RESULT 9  
AX357483  
LOCUS AX357483 800 bp DNA linear PAT 13-FEB-2002  
DEFINITION Sequence 21 from Patent WO0172773.  
ACCESSION AX357483  
VERSION AX357483.1 GI:18674539  
KEYWORDS .  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1  
AUTHORS Werner,G., Phares,W., Jaritz,M., Lapp,H. and Kalthoff,F.S.  
TITLE Organic compounds  
JOURNAL Patent: WO 0172773-A 21 04-OCT-2001;  
Novartis AG (CH)

FEATURES  
source Location/Qualifiers  
1..800  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 231 a 160 c 183 g 226 t

ORIGIN



[illegible]

LOCUS	AX456975	444 bp	DNA	linear	PAT 06-JUL-2002
DEFINITION	Sequence 14 from Patent WO0232958.				
ACCESSION	AX456975				
VERSION	AX456975.1	GI:21715770			
KEYWORDS	human.				
SOURCE	ORGANISM	Human sapiens			
REFERENCE	AUTHORS	Okuyama; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
TITLE	Polynucleotide and polypeptide sequences of human dendritic cell immunoreceptors				
JOURNAL	Patent: WO 0232958-A 14 25-APR-2002;				
FEATURES	Location/Qualifiers				
source	1..444	/organism="Homo sapiens"			
BASE COUNT	123 a 86 c 101 g 134 t				
ORIGIN					
Query Match	32.1%	Score 201.2;	DB 6;	Length 444;	
Best Local Similarity	68.5%	Pred. No. 9.4e-128;			
Matches	278;	Conservative 0;	Mismatches 128;	Indels 0;	Gaps 0;
Qy	222	AGCCTGGGGATGTTGCCCAGCTTCTTGGAAAGTCATTTGGTTCAGTTCGCTACTTCATTC 281			
Db	39	AGATTGGAGCTGCTGCCCAACCCCTTGGACTTCATTTTCAGTCTAGTTCGCTACTTCATTC 98			
Qy	282	CAGTCAAGAGAAAGTTTGGTCTAAGAGTGAGCAGAACTGTTGGAGTGGGAGCACATTT 341			
Db	99	TACTGGGATGCAATCTTGGACTTAAGAGTCAAAAGAACTGTTCTGTGATGGGGCTGATCT 158			
Qy	342	GTTGTGTTCAACACAGAGCAGACAAATTTTCATTCCTCAGCAGCTGAATGATCATT 401			
Db	159	GGTGGTGATCAACACACAGGGAAGAACAGGATTTTCATCATTCAGAACTCTGAAAGAAATTC 218			
Qy	402	TTCTATTATTTCTGGGGCTTTCAGAGCCCAACAGGTAATAATTAATTTGGCAATGGATTGATAA 461			
Db	219	TTCTATTATTTCTGGGGCTGTCAGATCCAGGGGTCGCGACATTTGGCAATGGTTGACCA 278			
Qy	462	GACACCTTATGAGAAAAATGTCAGATTTTGGCACCCTAGGTGAGGCCCAATTCATTCGCGAGA 521			
Db	279	GACACCACATAATGAAAAATGTACATTTCTGGCACTCAGGTGCAAGCCCAATAACCTTGATGA 338			
Qy	522	GCAATGTCCTTCAATAGTCCTTCTGGAACCTTACAGGATGGGCTGGAATGATGTTATCTG 581			
Db	339	CGTTGTGCGATAATAAATTTCCGTTCTTCAGAGAAATGGGGCTGGAATGACATTCACCTG 398			
Qy	582	TGAAACTAGAGGGAATTCATATGTGAGATGAATGAATGATTTACCTA 627			
Db	399	TCATGTACCTCAGAAGTCAATTTGCAAGATGAAGAGATCTACATA 444			
RESULT 13					
LOCUS	AX456962	402 bp	DNA	linear	PAT 06-JUL-2002
DEFINITION	Sequence 1 from Patent WO0232958.				
ACCESSION	AX456962				
VERSION	AX456962.1	GI:21715759			
KEYWORDS	human.				
SOURCE	ORGANISM	Human sapiens			
REFERENCE	AUTHORS	Okuyama; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
TITLE	Polynucleotide and polypeptide sequences of human dendritic cell immunoreceptors				
JOURNAL	Patent: WO 0232958-A 1 25-APR-2002;				
FEATURES	Location/Qualifiers				



source		1. 402		/organism="Homo sapiens"		/db_xref="taxon:9606"		112 t	
BASE COUNT		114 a		80 c		96 g		112 t	
ORIGIN									
Query Match		32.0%;		Score 200.4;		DB 6;		Length 402;	
Best Local Similarity		68.7%;		Pred. No. 1.6e-48;					
Matches		276;		Conservative		0;		Mismatches 126; Indels 0; Gaps 0;	
QY	226	TGGGATGTTGCCAGCTCTTGGAAAGTCATTTGGTCCAGTTGCTACTCTCAATTCACGT	285						
Db	1	TGGAGCTGCTGCCAACCCCTGGACTTCATTTTCAGTCTAGTTCGTTATTTCTACT	60						
QY	286	GAAGAGAAGTTTGTGCTAAGAGTGAGCAGAACTGTGTGAGATGGGAGCACATTTGGTT	345						
Db	61	GGATGCAATCTTGGACTTAAGAGTCAAAAGAACTGTTCTGTGATGGGGCTGATCTGGTG	120						
QY	346	GTGTTCAACAGACAGCAGCAAGAAATTTCAATTCAGCAGCTGAATGAGTCATTTTCT	405						
Db	121	GTGATCAACACAGGGAAGAACAGGATTTTCATCTCAGATCTGAAAAGAAATTCCTCT	180						
QY	406	TATTTTCTGGGGCTTCAGACCCACAGGTAAATAATTTGGCAATGGATGATAAGACA	465						
Db	181	TATTTTCTGGGGCTGTCTAGATCCAGGGGTGGGCACATTTGGCAATGGTTGACGACA	240						
QY	466	CCTATGAGAAAATGTCAGATTTTGGCACCTAGGTAGTCCCAATCAATTCGACAGACAA	525						
Db	241	CCATCAATGAAATGTGCATCTTGGCACTCAGGTGAACCAATAACCTTGATGAGCGT	300						
QY	526	TGTGCTTCAATAGTCTTCTGGAAACCTACAGATGGGGCTGGAATGATGTTATCTGTAA	585						
Db	301	TGTGCGATAATAAATTTCCGTTCTTCAGAAGAATGGGGCTGGAATGACATTCACCTGTCAT	360						
QY	586	ACTAGAAGAAATTCATATGTCAGATGAATGAATTTTACCTA	627						
Db	361	GTACCTCAGAAGTCAATTTGCAAGATGAAGAAGATCTACATA	402						
RESULT 14									
AF109146		Homo sapiens C-type lectin superfamily 6 (CLECSF6), mRNA		PRI 24-OCT-2000					
LOCUS		Homo sapiens C-type lectin superfamily 6 (CLECSF6), mRNA, complete cds.							
DEFINITION									
ACCESSION		AF109146		1133 bp		linear			
VERSION		AF109146.1		GI:6502534					
KEYWORDS		Homo sapiens.							
SOURCE		Homo sapiens							
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.							
REFERENCE		1 (bases 1 to 1133)							
AUTHORS		Richard M. and Beaulieu, A.D.							
TITLE		A novel C-type lectin expressed in GM-CSF stimulated neutrophils							
JOURNAL		Unpublished							
REFERENCE		2 (bases 1 to 1133)							
AUTHORS		Richard M. and Beaulieu, A.D.							
TITLE		Direct Submission							
JOURNAL		Submitted (24-NOV-1998) Dept. of Medicine, Centre Hospitalier de l'Universite Laval, 2705 Boul. Laurier, Sainte-Foy, Quebec G1V 4G2, Canada							
FEATURES		Location/Qualifiers							
source		1. .1133							
		/organism="Homo sapiens"							
		/db_xref="taxon:9606"							
gene		/cell_type="polymorphonuclear cells"							
1. .1133		/gene="CLECSF6"							
95. .808		/note="C-type lectin DB27"							
CDS		/codon_start=1							
		/product="C-type lectin superfamily 6"							

protein_id="AAFI4348.1"		/db_xref="GI:6502535"		/translation="MTSEITYAEVRFKNEFKSSGINTASSAASKERTAPLKSNTGPKK		LSCASLIIFFLLAIFVIFKYSQLEKTKTKELVHTTLCVKKMMPVEETA		WSCCPKNKWSFSSNCYFISTESASQDSEKDCARMEAHLLVINTQEQDFIFQNLQEE		SAYFVGLSDPEGORHWOVDQTPYNESSTFWHPRESDPNERCVLNFRKSPKRWGN		DVNCLGPQRSVCENMKIHL"		BASE COUNT		332 a		214 c		239 g		348 t	
ORIGIN																							
Query Match		30.6%;		Score 192;		DB 9;		Length 1133;															
Best Local Similarity		68.0%;		Pred. No. 5.5e-46;																			
Matches		283;		Conservative		0;		Mismatches 130;		Indels		3;		Gaps		1;							
QY	215	AGGTGCCAGCCTGGGGATGTTGCCCAGGCTTCTTGGAAAGTCATTTGGTGTCCAGTTGCTACT	274																				
Db	390	AAGAGACAGCCTGGAGCTGTTGCCCAAGAATTTGGAAGTCATTTAGTTTCCAACTGCTACT	449																				
QY	275	TCAATTTCCAGTCAAGAGAGGTTTGGTCTAAGAGTGAGCAGCAACTGTCTCAGATGGGAG	334																				
Db	450	TATTTTCTACTCAATCAGCATCTTGGCAAGACAGTGAGAAGGACTGTGCTAGAATGGAGG	509																				
QY	335	CACATTTGGTGTGTTCACACAGAGCAGACAGAAATTTCAATTTGCCAGCAGCTGAATG	394																				
Db	510	CTACCTGCTGGTGATAAACACTCAAGAAGAGCAGAGATTTTCATCTTCAGAAATCTGCAAG	569																				
QY	395	AGTCATTTTCTTATTTCTGGGGCTTTCAGACCCCAAGAGTAATAAATTTGGCAATGGA	454																				
Db	570	AAGAATCTGCTTATTTTGTGGGGCTCTCAGATCCAGAAAGGTCAGGACATTTGGCAATGGG	629																				
QY	455	TTCATAAGACACCTTATGAGAAAAATGTCAGATTTTGGCACCTAGGTGAGCCCAATCATTT	514																				
Db	630	TTGATCAGACACATACATGAAAGTTCCACATTTCCGATCCAGCTGAGCCCAAGTATC	689																				
QY	515	CTGCAGACGAATGTGCTTCAATAGTCTTCTGGAA---ACCTACAGGATGGGCTGGAATG	571																				
Db	690	CCAATGAGCGCTGCGTTGTGCTAAATTTTCGTAATAATCAACCAAGATGGGCTGGGAATG	749																				
QY	572	ATGTTATCTGTGAAACTAGAGAAGTAATCAATATGTGAGATGAATAAGATTTTACCTA	627																				
Db	750	ATGTTAATTTCTTGCTCTCCTCAAGGTCAGTTTGTGAGATGATGAAGATCCACATTA	805																				
RESULT 15																							
AX400064		Homo sapiens		1195 bp		DNA		linear		PAT 06-JUN-2002													
LOCUS		Sequence 235 from Patent WO218424.																					
DEFINITION																							
ACCESSION		AX400064		GI:21336479																			
VERSION		AX400064.1																					
KEYWORDS		human.																					
SOURCE		Homo sapiens																					
ORGANISM		Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.																					
REFERENCE		1																					
AUTHORS		Tang, Y.T., Asundi, V., Zhou, P., Xue, A.J., Ren, F., Zhang, J., Wang, J.R., Zhao, Q.A., Wang, D., Liu, C., Drmanac, R.T. and Wehrman, T.																					
TITLE		Nucleic acids and polypeptides																					
JOURNAL		Patent: WO 0218424-A 235 07-MAR-2002;																					
FEATURES		HYSEQ, INC. (US)																					
source		Location/Qualifiers																					
1. .1195		/organism="Homo sapiens"																					
CDS		/db_xref="taxon:9606"																					
270. .866		/note="unnamed protein product"																					
		/codon_start=1																					
		/db_xref="GI:21336480"																					
		/translation="MTSEITYAEVRFKNEFKSSGINTASSAASKERTAPLKSNTGPKK																					
		LVHTTLCVKKMMPVEETAWSCCPKNKSFSNNCYFISTESASQDSEKDCARMEAHLLVINTQEQDFIFQNLQEE																					
		SAYFVGLSDPEGORHWOVDQTPYNESSTFWHPRESDPNERCVLNFRKSPKRWGN																					
		DVNCLGPQRSVCENMKIHL"																					

BASE COUNT		340	a	224	c	268	g	363	t		
ORIGIN											
Query Match		30.6%; Score 192; DB 6; Length 1195;									
Best Local Similarity		68.0%; Pred. No. 5.5e-46;									
Matches 283; Conservative 0; Mismatches 130; Indels 3; Gaps 1;											
QY	215	AGGTGCCAGCCTGGGGATGTTGCCACAGCTTC	TTGGAAAGTCATTTGGTTCACAGTTGCTACT	274							
Db	448	AAGAGACAGCCTGGAGCTGTTGCCCAAGAAT	TGGAAGTCATTTAGTTCCAACTGCTACT	507							
QY	275	TCATTTCCAGTGAAGAGAAGGTTTGGTCTAAGAGT	GCAGCAGAACTGCTGTGAGATGGGAG	334							
Db	508	TTATTTCTACTGAATCAGCATCTTGGCAAGACAGT	GCAGAGGACTGCTGTAGAAATGGAGG	567							
QY	335	CACATTTGGTTGTGTTCAACACAGACAGCAGCA	AAATTTTCATTTGCCAGCAGCTGAATG	394							
Db	568	CTCACCTGCTGGTGATAAACACTCAAGAAGCAGG	ATTTTCATCTTCCAGAAATCTGCAAG	627							
QY	395	AGTCATTTTCTTATTTTCTGGGGCTTTTCAGACCC	ACAGGTAATAATATTTGGCAATGGA	454							
Db	628	AAGAATCTGCTTATTTTGTGGGGCTCTCAGATCC	AGAGGTCAGCGACATTTGGCAATGGG	687							
QY	455	TTGATAAGACACCTTATCAGAAAAATGTCAGAT	TTTGGCACCTAGGTGAGCCCAATCATT	514							
Db	688	TTGATCAGACACCATACAAATGAAAGTTCCAT	ATTTCTGGCATCCACGTGAGCCCACTGATC	747							
QY	515	CTGCAGAGCAATGTGCTTCAATAGTCTTCTGGAA	---ACCTACAGGATGGGGCTGGAATG	571							
Db	748	CCAATGAGCGCTGCGTTGTGCTAAATTTTCGT	AAATCACCCAAAAGATGGGCTGGAATG	807							
QY	572	ATGTTATCTGTGAAGCTAGAAGGAATTCATAT	TGTGAGATGAATAGATTTACCTA	627							
Db	808	ATGTTAAATGTCTTGGTCTCCTCAAAGGTCAG	TTTGTGAGATGATGAAGATCCCACTTA	863							

Search completed: February 17, 2003, 11:03:47  
Job time : 1085.37 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: February 20, 2003, 19:28:17 ; Search time 47 Seconds  
(without alignments)  
1363.733 Million cell updates/sec

Title: US-09-766-511B-53  
Perfect score: 1165  
Sequence: 1 MMQEQPQSTKRGWLSRL.....NDVICETRRNSICEMNKIYL 209

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US09766511/runat\_19022003\_094339\_29316/app\_query.fasta\_1.391  
-DB=Issued Patents\_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -NORM=ext -HEAPSIZE=500 -MINLEN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09766511@cgn\_1.1.17/runat\_19022003\_094339\_29316 -NCFU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents\_NA:\*

1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*

2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*

3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*

4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*

5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq:\*

6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	793	68.1	1227	3	US-08-772-440-3
2	679	58.3	501	3	Sequence 3, Appli
3	565	48.5	393	3	Sequence 20, Appl
4	458	39.3	1104	4	Sequence 22, Appl
5	436	37.4	1418	4	Sequence 1, Appli
6	251.5	21.6	1458	4	Sequence 7, Appli
7	250	21.5	1370	4	Sequence 9, Appli
8	235	20.2	1212	4	Sequence 3, Appli
9	235	20.2	1212	4	Sequence 9, Appli
10	234	20.1	1212	4	Sequence 1, Appli
11	232	19.9	1212	4	Sequence 11, Appl
12	229.5	19.7	10409	3	Sequence 10, Appl
					Sequence 33, Appli

13	223	19.1	152	3	US-08-772-440-40	Sequence 40, Appl
14	192.5	16.4	4771	3	US-08-840-062-3	Sequence 3, Appli
15	191	16.4	145	3	US-08-772-440-42	Sequence 42, Appl
16	189.5	16.3	4588	3	US-08-840-062-1	Sequence 1, Appli
17	187	16.1	1897	2	US-08-809-494A-1	Sequence 1, Appli
18	187	16.1	1897	4	US-09-352-302-1	Sequence 1, Appli
19	182.5	15.7	3259	5	PCT-US95-03747-1	Sequence 1, Appli
20	180.5	15.5	885	1	US-08-365-103B-3	Sequence 3, Appli
21	180.5	15.5	924	1	US-08-365-103B-5	Sequence 5, Appli
22	180.5	15.5	1005	1	US-08-365-103B-1	Sequence 1, Appli
23	179	15.4	1306	2	US-08-809-494A-3	Sequence 3, Appli
24	179	15.4	1306	2	US-09-352-302-3	Sequence 3, Appli
25	172	14.8	116	3	US-08-772-440-41	Sequence 41, Appl
26	169.5	14.5	561	4	US-09-535-521-16	Sequence 16, Appl
27	169.5	14.5	561	4	US-09-535-521-18	Sequence 18, Appl
28	169.5	14.5	624	4	US-09-535-521-19	Sequence 19, Appl
29	169.5	14.5	624	4	US-09-535-521-21	Sequence 21, Appl
30	169.5	14.5	876	4	US-09-535-521-4	Sequence 4, Appli
31	169.5	14.5	876	4	US-09-535-521-6	Sequence 6, Appli
32	169.5	14.5	2851	4	US-09-535-521-1	Sequence 1, Appli
33	169.5	14.5	2851	4	US-09-535-521-3	Sequence 3, Appli
34	168	14.4	996	4	US-09-127-946-11	Sequence 11, Appl
35	167.5	14.4	1461	3	US-08-722-126A-4	Sequence 4, Appli
36	167.5	14.4	1461	5	PCT-US95-04258-4	Sequence 10, Appl
37	166.5	14.3	417	4	US-09-535-521-10	Sequence 12, Appl
38	166.5	14.3	417	4	US-09-535-521-12	Sequence 13, Appl
39	166.5	14.3	423	4	US-09-535-521-13	Sequence 15, Appl
40	166.5	14.3	423	4	US-09-535-521-15	Sequence 7, Appli
41	166	14.2	384	4	US-09-535-521-7	Sequence 9, Appli
42	166	14.2	384	4	US-09-535-521-9	Sequence 2, Appli
43	166	14.2	5169	4	US-09-194-612A-2	Sequence 24, Appl
44	165.5	14.2	369	4	US-09-535-521-24	Sequence 2, Appli
45	165.5	14.2	369	4	US-09-535-521-26	Sequence 26, Appl

#### ALIGNMENTS

RESULT 1

US-08-772-440-3

; Sequence 3, Application US/08772440

; Patent No. 6046158

; GENERAL INFORMATION:

; APPLICANT: Arizumi, Kiyoshi

; APPLICANT: Takashima, Akira

; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE

; TITLE OF INVENTION: LECTINS, DECTIN-1 AND DECTIN-2; COMPOSITIONS AND USES

; TITLE OF INVENTION: THEREOF

; NUMBER OF SEQUENCES: 42

; CORRESPONDENCE ADDRESS:

; ADDRESS: Arnold, White & Durkee

; STREET: P.O. Box 4433

; CITY: Houston

; STATE: Texas

; COUNTRY: USA

; ZIP: 77210

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/772.440

; FILING DATE: CONCURRENTLY HERewith

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Parker, David L.

; REGISTRATION NUMBER: 32,165

; REFERENCE/DOCKET NUMBER: UTXD.493

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 512/418-3000

; TELEFAX: 512/474-7577

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:



Db	481	GAATGAAGAGATTACCTA	501
Qy	179	IleValpheTrpLysProThrGlyTrpAsnAspValIleCysGluThrArgArg	199
Db	301	ATAGTTTACTGGAATCCCTTCGAAATGGGGCTGGAATGATGTTTCTGTCATAGTAAACAC	360
Qy	199	AsnSerIleCysGluMetAsnLysIleTyrLeu	209
Db	361	AATTCAATATGTGAATGAAGAAGATTACCTA	393
<p>RESULT 4</p> <p>US-09-111-470-1</p> <p>; Sequence 1, Application US/09111470</p> <p>; Patent No. 6277959</p> <p>; GENERAL INFORMATION:</p> <p>; APPLICANT: Valladeau, Jenny</p> <p>; APPLICANT: Ravel, Odile</p> <p>; APPLICANT: Bates, Elizabeth E.M.</p> <p>; APPLICANT: Ford, John</p> <p>; APPLICANT: Saeland, Sem</p> <p>; APPLICANT: Lebecque, Serge J.E.</p> <p>; TITLE OF INVENTION: Mammalian Membrane Protein Genes;</p> <p>; TITLE OF INVENTION: Related Reagents</p> <p>; NUMBER OF SEQUENCES: 11</p> <p>; CORRESPONDENCE ADDRESS:</p> <p>; ADDRESSEE: DNAX Research Institute</p> <p>; STREET: 901 California Avenue</p> <p>; CITY: Palo Alto</p> <p>; STATE: California</p> <p>; COUNTRY: USA</p> <p>; ZIP: 94304-1104</p> <p>; COMPUTER READABLE FORM:</p> <p>; MEDIUM TYPE: Floppy disk</p> <p>; OPERATING SYSTEM: PC-DOS/MS-DOS</p> <p>; SOFTWARE: PatentIn Release #1.0, Version #1.30</p> <p>; CURRENT APPLICATION DATA:</p> <p>; APPLICATION NUMBER: US/09/111,470</p> <p>; FILING DATE: 08-JUL-1998</p> <p>; CLASSIFICATION:</p> <p>; PRIOR APPLICATION DATA:</p> <p>; APPLICATION NUMBER: US 60/053,080</p> <p>; FILING DATE: 09-JUL-1997</p> <p>; ATTORNEY/AGENT INFORMATION:</p> <p>; NAME: Ching, Edwin P.</p> <p>; REGISTRATION NUMBER: 34,090</p> <p>; REFERENCE/DOCKET NUMBER: SF0695</p> <p>; TELECOMMUNICATION INFORMATION:</p> <p>; TELEPHONE: (650)852-9196</p> <p>; TELEFAX: (650)496-1200</p> <p>; INFORMATION FOR SEQ ID NO: 1:</p> <p>; SEQUENCE CHARACTERISTICS:</p> <p>; LENGTH: 1104 base pairs</p> <p>; TYPE: nucleic acid</p> <p>; STRANDEDNESS: single</p> <p>; TOPOLOGY: linear</p> <p>; MOLECULE TYPE: cDNA</p> <p>; FEATURE:</p> <p>; NAME/KEY: CDS</p> <p>; LOCATION: 242..952</p> <p>US-09-111-470-1</p>			
<p>Alignment Scores:</p> <p>Pred. No.: 2,82e-44 Length: 1104</p> <p>Score: 458.00 Matches: 83</p> <p>Percent Similarity: 62.57% Conservative: 34</p> <p>Best Local Similarity: 44.39% Mismatches: 66</p> <p>Query Match: 39.31% Indels: 4</p> <p>DB: Gaps: 3</p>			
<p>US-09-766-511B-53 (1-209) x US-09-111-470-1 (1-1104)</p>			
Qy	26	IleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysValValThrTyrHisPhe	45
Db	241	CAAAATGTCAGGTTCTGGCACCCCATGAACCAATCTTCCAGAAAGCGGTGTGTTCA	300



Db 954 AACAGAGTCTGTTCTCAGATGAGAGAAATAAACTTA 992

RESULT 6

US-09-111-470-3

; Sequence 3, Application US/09111470

; Patent No. 6277959

; GENERAL INFORMATION:

; APPLICANT: Valladeau, Jenny

; APPLICANT: Ravel, Odile

; APPLICANT: Bates, Elizabeth E.M.

; APPLICANT: Ford, John

; APPLICANT: Saeland, Sem

; APPLICANT: Lebecque, Serge J.E.

; TITLE OF INVENTION: Mammalian Membrane Protein Genes;

; TITLE OF INVENTION: Related Reagents

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: DNAX Research Institute

; STREET: 901 California Avenue

; CITY: Palo Alto

; STATE: California

; COUNTRY: USA

; ZIP: 94304-1104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/111,470

; FILING DATE: 08-JUL-1998

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/053,080

; FILING DATE: 09-JUL-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Ching, Edwin P.

; REGISTRATION NUMBER: 34,090

; REFERENCE/DOCKET NUMBER: SF0695

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (650)852-9196

; TELEFAX: (650)496-1200

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1458 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 257..1204

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: 608

; OTHER INFORMATION: /note= "short form lacks

; OTHER INFORMATION: nucleotides 608-673"

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: 775

; OTHER INFORMATION: /note= "ASGPRm (table 2) has

; OTHER INFORMATION: sequence insert encoding GEE between nucleotides 775-776"

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: 1064

; OTHER INFORMATION: /note= "nucleotide 1064 of DCMp2s

; OTHER INFORMATION: may be A, which would encode asn rather than asp at the resid

; OTHER INFORMATION: numbered 270"

; US-09-111-470-3

Alignment Scores:

Pred. No.: 7,66e-20 Length: 1458

Score: 251.50 Matches: 55

Percent Similarity: 49.70% Conservative: 28  
Best Local Similarity: 32.93% Mismatches: 55  
Query Match: 21.59% Indels: 29  
DB: 4 Gaps: 7

US-09-766-511b-53 (1-209) x US-09-111-470-3 (1-1458)

QY 53 ArgLeuSerGluLeuHisSerTyrHisSerSerLeuThrCys----- 66  
Db 710 CGAGTCCAGCAGCTGGTGCAGAACCTGGAAGAACTGACCTGCCAGGTGGCTACTCTCAAC 769  
QY 67 -----PheSerGluGlyThrLysValProAlaTrpGlyCysCysProAlaSerTrp 83  
Db 770 AACAAATGCTCCAGTGAAGGACC-----TGCTGCCCTCACTGAG 811  
QY 84 LysSerPheGlySerSerCysTyrPheLeSerSerGluGluLysValTrpSerLysSer 103  
Db 812 GTGGAGCACCAAGACAGCTGCTACTGTTCTCTCCTCTGGGATGCTCTGGCGGAGGCT 871  
QY 104 GluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsnThrGluAlaGluGln 123  
Db 872 GAGAGTACTGCTGAGTGAAGAACGCCACCTGGTGGTCACTCACTCCAGGAGGAGGAG 931  
QY 124 AsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeuSerAspPro 143  
Db 932 AATTTTGTCCAGAAATATCTAGGCTCCGATACACCTGG---ATGGCCTCAGTGACCT 988  
QY 144 GlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsnValArgPhe 163  
Db 989 GAAGGA-----GCCTGGAAGTGGGTGATGGAACAGACTATCGACCGGCTTCCAGAAC 1042  
QY 164 TrpHisLeuGlyGluProAsnHis-----SerAlaGluGlnCys 176  
Db 1043 TGAAGCCAGGCCAGCCAGCCAGCTGGAGGGGACGGGTGGGTGGAGGCGGAGCTGT 1102  
QY 177 AlaSerIleValPheTrpLysProThrGlyTrpGlyTrpGlyTrpAsnAspValIleCysGluThr 196  
Db 1103 GCT-----CACTTCCATCCAGACGCGC---AGGTGGAATGACGACGCTCTGCCAGAG 1150  
QY 197 ArgArgAsnSerIleCysGlu 203  
Db 1151 CCCTACCACCTGGGTCTGCGAG 1171

RESULT 7

US-09-111-470-9

; Sequence 9, Application US/09111470

; Patent No. 6277959

; GENERAL INFORMATION:

; APPLICANT: Valladeau, Jenny

; APPLICANT: Ravel, Odile

; APPLICANT: Bates, Elizabeth E.M.

; APPLICANT: Ford, John

; APPLICANT: Saeland, Sem

; APPLICANT: Lebecque, Serge J.E.

; TITLE OF INVENTION: Mammalian Membrane Protein Genes;

; TITLE OF INVENTION: Related Reagents

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: DNAX Research Institute

; STREET: 901 California Avenue

; CITY: Palo Alto

; STATE: California

; COUNTRY: USA

; ZIP: 94304-1104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/111,470

; FILING DATE: 08-JUL-1998

; CLASSIFICATION:

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/053,080
; FILING DATE: 09-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: SF0695
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)852-9196
; TELEFAX: (650)496-1200
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1370 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 273..1091
US-09-111-470-9

Alignment Scores:
Pred. No.: 1,04e-19 Length: 1370
Score: 250.00 Matches: 55
Percent Similarity: 48.82% Conservative: 28
Best Local Similarity: 32.35% Mismatches: 55
Query Match: 21.46% Indels: 32
DB: 4 Gaps: 7

US-09-766-511B-53 (1-209) x US-09-111-470-9 (1-1370)
QY 53 ArgLeuSerGluHisSerTyrHisSerSerLeuThrCys----- 66
Db 588 CGAGTCCAGCAGCTGGTCCAGACCTGAGAACTGACCTGCCAGGTGGTACTCTCAAC 647
QY 67 -----PheSerGluGlyThrLysValProAlaTrpGlyCysCysPro 80
Db 648 AACAATGGTGGAGAACCTCCACTCAAGGACC-----TGCTGCCCC 689
QY 81 AlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGluLysValTrp 100
Db 690 GTCAACTGGTGGACACACAGACAGCTGCTACTGGTTCTCTCACTCGGATGTCCTGG 749
QY 101 SerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsnThrGlu 120
Db 750 GCCAGGCTGAGAAGTACTGCCAGTGAAGAACGCCACCTGGTGGTCATCACTCCAGG 809
QY 121 AlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeu 140
Db 810 GAGGACACAGAAATTTGTCCAGAAATATCTAGGCTCCGCATACAGCTGG---ATGGGCTC 866
QY 141 SerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsn 160
Db 867 AGTGACCTGTAAGGA-----GCCTGGAAGTGGGTGGATGAACAGACTATGCCACCGCC 920
QY 161 ValArgPheTrpHisLeuGlyGluProAsnHis-----SerAla 173
Db 921 TTCAGAACTGGAAGCCAGCCAGCAGCAGCTGGCAGGGCCAGCGGTGGTGAGGC 980
QY 174 GluGlnCysAlaSerIleValPheTrpLysProThrGlyTrpAsnAspValIle 193
Db 981 GAGGACTGTGCT-----CACTTCCATCCAGACGCG---AGTGGATGACGACGTC 1028
QY 194 CysGluThrArgArgAsnSerIleCysGlu 203
Db 1029 TCCAGAGGCCCTACCACCTGGGTCTGCCAG 1058

RESULT 8
US-09-591-435-9
; Sequence 9, Application US/09591435
; Patent No. 6280953
; GENERAL INFORMATION:
; APPLICANT: MESSIER, WALTER
; APPLICANT: SIKELA, JAMES M
; TITLE OF INVENTION: METHODS TO IDENTIFY POLYNUCLEOTIDE AND POLYPEPTIDE
; TITLE OF INVENTION: SEQUENCES WHICH MAY BE ASSOCIATED WITH PHYSIOLOGICAL
; TITLE OF INVENTION: AND MEDICAL CONDITIONS
; FILE REFERENCE: GENO.200.2
; CURRENT APPLICATION NUMBER: US/09/591,435
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 09/591,435
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 09/240,915
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: 60/073,263
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/098,987
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-591-435-9

Alignment Scores:
Pred. No.: 4.95e-18 Length: 1212
Score: 235.00 Matches: 48
Percent Similarity: 55.47% Conservative: 23
Best Local Similarity: 37.50% Mismatches: 49
Query Match: 20.17% Indels: 8
DB: 4 Gaps: 4

US-09-766-511B-53 (1-209) x US-09-591-435-9 (1-1212)
QY 79 CysProAlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGluGluLys 98
Db 766 TGTCCCTGGGAATGGACATCTTCCAGGAAGTCTTACTTCTAATCTCCAGCGG 825
QY 99 ValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsn 118
Db 826 AACTGGCAGCAGCTCCATCATCCGCGCTGCAAGAAGTGGGGCCAGCTCGCTAATCAAA 885
QY 119 ThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyr---Phe 137
Db 886 AGTGCTGAGGAGACAGAACTCTCTACAGCTGCTTCCAGAAAGTAACCGCTTCACTGG 945
QY 138 LeuGlyLeuSerAspProGlnGlnGlyAsnAsnTrpGlnTrpIleAspLysThrProTyr 157
Db 946 ATGGGACTTTCAGATCTAATATCAGGAAGGCACCTGGCAATGGGTGGACGCTCACCTCTG 1005
QY 158 GluLysAsnValArg---PheTrpHisLeuGlyGluProAsnHisSerAlaGluGln--- 175
Db 1006 TTGCCCAGCTTCAAGCAGTATTGGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1065
QY 176 CysAlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGlu 195
Db 1066 TGGCGCGGAATTT-----AGTGGCAATGGCTGGAAACGACGACAAATGTAAT 1110
QY 196 ThrArgArgAsnSerIleCysGlu 203
Db 1111 CTTGCCCAAAATCTGGATCTGCAAA 1134

RESULT 9
US-09-517-605-1
; Sequence 1, Application US/09517605
; Patent No. 6391567
; GENERAL INFORMATION:
; APPLICANT: Littman, Dan R.
; APPLICANT: Kwon, Douglas S.
; APPLICANT: van Kooyk, Yvette
; APPLICANT: Geijtenbeck, Theo
; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO
; TITLE OF INVENTION: CELLS
```



```
; FILE REFERENCE: 1049-1-017
; CURRENT APPLICATION NUMBER: US/09/517,605
; CURRENT FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1312
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (42)..(1253)
US-09-517-605-1
Alignment Scores:
Pred. No.: 5,61e-18 Length: 1312
Score: 235.00 Matches: 48
Percent Similarity: 55.47% Conservative: 23
Best Local Similarity: 37.50% Mismatches: 49
Query Match: 20.17% Indels: 8
DB: 4 Gaps: 4

US-09-766-511B-53 (1-209) x US-09-517-605-1 (1-1312)
QY 79 CysProAlaSerTrpLysSerPheGlySerCysTyrPheIleSerSerGluGluLys 98
Db 807 TGTCCCTGGGAATGGACATCTTCCAAAGAACTGTACTTCATGCTAACTCCAGCGG 866
QY 99 valTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsn 118
Db 867 AACTGGCAGCAGCTCCATCAGCGCTGCCAAAGAGTGGGGGCCAGCGTCGCTGAATCAAA 926
QY 119 ThrGluAlaGlnAsnPheIleValGlnLeuAsnGluSerPheSerTyr---Phe 137
Db 927 AGTCTGAGGAGCAGAACTCTTACACCTGCAGTCTTCCAGAGATTAACCGCTTACACCTGG 986
QY 138 LeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrProTyr 157
Db 987 ATGGGACTTTCAGATCTAAATCAGGAAGCAGCGTGGCAATGGGTGGACGGCTCACCTCTG 1046
QY 158 GluLysAsnValArg---PheTrpHisLeuGlyGluProAsnHisSerAlaGluGln--- 175
Db 1047 TTGCCCAGAGCTTCAAGCAGTATTGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1106
QY 176 CysAlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGlu 195
Db 1107 TGCCTGGGAATTT-----AGTGGCAATGGCTGGGAACGACGACAAATGTAAT 1151
QY 196 ThrArgArgAsnSerIleCysGlu 203
Db 1152 CTTGCCAAATCTGGATCTGCAAA 1175

RESULT 10
US-09-591-435-11
; Sequence 11, Application US/09591435
; Patent No. 6280953
; GENERAL INFORMATION:
; APPLICANT: MESSIER, WALTER
; APPLICANT: SIKELA, JAMES M
; TITLE OF INVENTION: METHODS TO IDENTIFY POLYNUCLEOTIDE AND POLYPEPTIDE
; TITLE OF INVENTION: SEQUENCES WHICH MAY BE ASSOCIATED WITH PHYSIOLOGICAL
; TITLE OF INVENTION: AND MEDICAL CONDITIONS
; FILE REFERENCE: GENO.200.2
; CURRENT APPLICATION NUMBER: US/09/591,435
; CURRENT FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 09/591,435
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 09/240,915
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: 60/073,263
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/098,987
; PRIOR FILING DATE: 1998-09-02

; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Pan troglodytes
```

```
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Gorilla gorilla
US-09-591-435-11
Alignment Scores:
Pred. No.: 6,49e-18 Length: 1212
Score: 234.00 Matches: 49
Percent Similarity: 54.96% Conservative: 23
Best Local Similarity: 37.40% Mismatches: 45
Query Match: 20.09% Indels: 14
DB: 4 Gaps: 5

US-09-766-511B-53 (1-209) x US-09-591-435-11 (1-1212)
QY 79 CysProAlaSerTrpLysSerPheGlySerCysTyrPheIleSerSerGluGluLys 98
Db 766 TGCCTCTGGGAATGGACATCTTCCAAAGAACTGTACTTCATGCTAACTCCAGCGG 825
QY 99 ValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsn 118
Db 826 AACTGGCAGCAGCTCCATCAGCGCTGCCAAAGAGTGGGGGCCAGCTCGTGTAAATCAAA 885
QY 119 ThrGluAlaGlnAsnPheIleValGlnLeuAsnGluSerPheSerTyr---Phe 137
Db 886 AGTCTGAGGAGCAGAACTTCTACAGCTGCAGTCTTCCAGAGAGTAACCGCTTACCTGG 945
QY 138 LeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrPro--- 156
Db 946 ATGGGACTTTCAGATCTAAATCATGAGCAGCTGGCAATGGGTGGACGGCTCACCTCTG 1005
QY 157 -----TyrGluLysAsnValArgPheTrpHisLeuGlyGluProAsnHisSerAla 173
Db 1006 TTGCCCAGCTTCGAG-----CAGTATTGGAACAGAGAGAGAGAGAGAGAGAGAGAG 1056
QY 174 GluGln---CysAlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspVal 192
Db 1057 GAGGAGAGCTGGCGGAATTT-----AGTGGCAATGGCTGGGAACGATGAC 1101
QY 193 IleCysGluThrArgArgAsnSerIleCysGlu 203
Db 1102 AAATGTAATCTTGCCAAATCTGGATCTGCAAA 1134

RESULT 11
US-09-591-435-10
; Sequence 10, Application US/09591435
; Patent No. 6280953
; GENERAL INFORMATION:
; APPLICANT: MESSIER, WALTER
; APPLICANT: SIKELA, JAMES M
; TITLE OF INVENTION: METHODS TO IDENTIFY POLYNUCLEOTIDE AND POLYPEPTIDE
; TITLE OF INVENTION: SEQUENCES WHICH MAY BE ASSOCIATED WITH PHYSIOLOGICAL
; TITLE OF INVENTION: AND MEDICAL CONDITIONS
; FILE REFERENCE: GENO.200.2
; CURRENT APPLICATION NUMBER: US/09/591,435
; CURRENT FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 09/591,435
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 09/240,915
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: 60/073,263
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/098,987
; PRIOR FILING DATE: 1998-09-02

; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Pan troglodytes
```

## US-09-591-435-10

## Alignment Scores:

Pred. No.: 1.11e-17 Length: 1212  
Score: 232.00 Matches: 48  
Percent Similarity: 55.47% Conservative: 23  
Best Local Similarity: 37.50% Mismatches: 49  
Query Match: 19.91% Indels: 8  
DB: 4 Gaps: 4

US-09-766-511b-53 (1-209) x US-09-591-435-10 (1-1212)

QY 79 CysProAlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGluGluLys 98  
||||| ||| ||| :|||:|||||:||||| :|||:  
Db 766 TGCCCTGGGAATGACATCTTCAAGAGAAATGTTACTTCATGCTAACTCCACGGG 825  
QY 99 ValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsn 118  
||| ||| ||| |||:||||| |||:|||||  
Db 826 AACTGGCAGCACTCCATCAGTCCTGCAAGAGAGTGGGGGCCAGCTCGTGTAAATCAAA 885  
QY 119 ThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyr---Phe 137  
||| |||:|||||: ||| :||| :||| :|||  
Db 886 AGTGTGAGGAGCAGAACTCTCTACAGCTGCAGTCTTCCAGAAGTAACCGCTTACCTGG 945  
QY 138 LeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrProTyr 157  
:|||||:||||| |||:|||||:||||| :|||  
Db 946 ATGGACTTTCAGATCTAAATGAGGAGGCATGTGGCAATGGTGGACGGCTCACCTCTG 1005  
QY 158 GluLysAsnVal---ArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGln--- 175  
:|: :|||:|||||: |||:|||||:||||| :|||:  
Db 1006 TTGCCCACTTCAACGAGTATGGAAACAGAGAGAGCCCAACAGTGGGGAGGAAGAC 1065  
QY 176 CysAlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGlu 195  
||||| ||| |||:||||| |||  
Db 1066 TCCGCGGAATT-----AGTGGCAATGCTGGAATGAGGACAAATGTAAT 1110  
QY 196 ThrArgArgAsnSerIleCysGlu 203  
:|: :|||:|||||:  
Db 1111 CTTGCCAAATCTGGATCGCAA 1134

## RESULT 12

US-08-772-440-33

; Sequence 33, Application US/08772440  
; Patent No. 6046158  
; GENERAL INFORMATION:  
; APPLICANT: Ariizumi, Kiyoshi  
; APPLICANT: Takashima, Akira  
; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE  
; TITLE OF INVENTION: LECTINS, DECTIN-1 AND DECTIN-2; COMPOSITIONS AND USES  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ARNOLD, White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/772.440  
; FILING DATE: CONCURRENTLY HERewith  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Parker, David L.  
; REGISTRATION NUMBER: 32,165  
; REFERENCE/DOCKET NUMBER: UTXD:493  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 512/418-3000

; TELEFAX: 512/474-7577  
; INFORMATION FOR SEQ ID NO: 33:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10409 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: 6510  
; OTHER INFORMATION: /mod\_base= OTHER  
; OTHER INFORMATION: /note= "D = A or G or T"  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: 3406..6470  
; OTHER INFORMATION: /mod\_base= OTHER  
; OTHER INFORMATION: /note= "K = G or T"  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: 3564..7896  
; OTHER INFORMATION: /mod\_base= OTHER  
; OTHER INFORMATION: /note= "M = A or C"  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: 3497..3607  
; OTHER INFORMATION: /mod\_base= OTHER  
; OTHER INFORMATION: /note= "N = A or C or G or T"  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: 3479..6422  
; OTHER INFORMATION: /mod\_base= OTHER  
; OTHER INFORMATION: /note= "R = A or G"  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: 3405..6871  
; OTHER INFORMATION: /mod\_base= OTHER  
; OTHER INFORMATION: /note= "S = C or G"  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: 3457..9998  
; OTHER INFORMATION: /mod\_base= OTHER  
; OTHER INFORMATION: /note= "W = A or T"  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: 3595..9999  
; OTHER INFORMATION: /mod\_base= OTHER  
; OTHER INFORMATION: /note= "Y = C or T"  
; US-08-772-440-33  
Alignment Scores:  
Pred. No.: 6.49e-16 Length: 10409  
Score: 229.50 Matches: 49  
Percent Similarity: 54.39% Conservative: 13  
Best Local Similarity: 42.98% Mismatches: 24  
Query Match: 19.70% Indels: 28  
DB: 3 Gaps: 3  
US-09-766-511b-53 (1-209) x US-08-772-440-33 (1-10409)  
QY 76 TrpGlyCysCysProAlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSer 95  
||||| ||| ||| :|||:|||||:||||| :|||:  
Db 5755 TGGGGATGCTGCCCAATCACTGGAAGTCAATTTGGGCTCCAGCTGCTACCTCATTTCTACC 5814  
QY 96 GluGluLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuVal 115  
:||||| |||:||||| :|||:  
Db 5815 AAGGAGAACTTCTGGACCACTGAGCAGAACTGTTCAGATGGGGCTCATCTGGTG 5874  
QY 116 ValPheAsnThrGluAlaGluGln----- 123  
||| |||:|||||: |||  
Db 5875 GTGATCAATACTGAAGCGGAGCAGTAGTCTTTCTTTTTCATTAGCCTTTTAATTG 5934  
QY 124 -----Asn-PheIleValGlnGlnLeuAsnGluSerPheSerTyrPh 137  
||| :||| :||| :||| :||| :|||

```
Db 5935 TAGGAAATGTAGTTTGAAATGTATGTTTCATCCCATCTAGTATGTAAGAAAGCCCTTAGAATTTT 5994
QY 137 e-----LeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAs 153
      |||
Db 5995 TATACCTAGGAGATTGGATTCTCTCTCCAGTGGTAGTCGAAT-----6040
      |||
QY 153 pLysThrProTyrGluLysAsnValArgPheTrpHisLeu 166
      |||
Db 6041 -AGACTACCCACTTCCCAATCCATAAGGTTCTGGCAATTA 6079
      |||
RESULT 13
US-08-772-440-40
; Sequence 40, Application US/08772440
; Patent No. 6046158
; GENERAL INFORMATION:
; APPLICANT: Ariizumi, Kiyoshi
; APPLICANT: Takashima, Akira
; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE
; TITLE OF INVENTION: LECTINS, DECTIN-1 AND DECTIN-2; COMPOSITIONS AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/772,440
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTXD:493
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 152 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-772-440-40
Alignment Scores:
Pred. No.: 4.82e-18 Length: 152
Score: 223.00 Matches: 38
Percent Similarity: 85.42% Conservative: 3
Best Local Similarity: 79.17% Mismatches: 7
Query Match: 19.14% Indels: 0
DB: Gaps: 0
US-09-766-511b-53 (1-209) x US-08-772-440-40 (1-152)
QY 76 TrpGlyCysProAlaSerTrpLysSerPheGlySerCysTyrPheIleSerSer 95
      |||
Db 9 TGGGGATGCTGCCCAATCACTGGAGTCAATTTGGCTCCAGCTGCATCTTCTACC 68
      |||
QY 96 GluGluLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuVal 115
      |||
Db 69 AAGGAGAACTTCTGGAGCACCAGTGAGCAGCACTGTGTTCCAGATGGGGGCTCATCTGGTG 128
      |||
QY 116 ValPheAsnThrGluAlaGluGln 123
      |||
Db 129 GTGATCACTACTGAAGCGGAGCAG 152
      |||
```

```
RESULT 14
US-08-840-062-3
; Sequence 3, Application US/08840062
; Patent No. 6117977
; GENERAL INFORMATION:
; APPLICANT: LASKY, LAURENCE A.
; APPLICANT: WU, KAI
; TITLE OF INVENTION: TYPE C LECTINS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/840,062
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P1019r1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3216
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4771 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-840-062-3
Alignment Scores:
Pred. No.: 4.19e-12 Length: 4771
Score: 192.50 Matches: 43
Percent Similarity: 52.71% Conservative: 25
Best Local Similarity: 33.33% Mismatches: 52
Query Match: 16.52% Indels: 9
DB: Gaps: 5
US-09-766-511b-53 (1-209) x US-08-840-062-3 (1-4771)
QY 79 CysProAlaSerTrpLysSerPheGlySerCysTyrPheIleSerSerGluGluLys 98
      |||
Db 1156 TCGAGCGCGAGCTGGCGAGCCCTTCAGGGCCACTGTACCGCTGCAGGCCGAGAAGCGC 1215
      |||
QY 99 ValTrpSerLysSerClnAsnCysValGluMetGlyAlaHisLeuValPheAsn 118
      |||
Db 1216 AGCTGGCAGGAGTCCAAAGAGGCATGCTACGGGGGGTGGCGACCTGGTCAGCATCCAC 1275
      |||
QY 119 ThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPhe---SerTyrPhe 137
      |||
Db 1276 AGCATGGCGGAGCTGGAATTCATCACCAGCAGATCAACAGAGGTGGAGGAGCTGTGG 1335
      |||
QY 138 LeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrProTyr 157
      |||
Db 1336 ATCGGCTCAACGATTTGAAGCTGCAGATGAATTTTGTAGTGTCTCAGCGGAGCTTGTG 1395
      |||
QY 158 GluLysAsnValArgPheTrpHisLeuGlyGluProAsnHis-----SerAlaGlu 174
      |||
Db 1396 -----AGCTTCACCCACTGGGACCCCTTTGAGCCCAACAACTTCGGGACAGCTGGAG 1449
      |||
QY 175 GlnCysAlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCys 194
      |||
```

Job time : 52 secs

```

Db 1450 GACTGTGTACCATC-----TGGGGCCCGGAAGGC---CGCTGGAACGACAGTCCCTGT 1500
QY 195 GluThrArgAsnSerIleCysGlu 203
Db 1501 AACCACTCTTGCATCCATCCATCTGCAAG 1527

RESULT 15
US-08-772-440-42
; Sequence 42, Application US/08772440
; Patent No. 6046158
; GENERAL INFORMATION:
; APPLICANT: Arizumi, Kiyoshi
; APPLICANT: Takashima, Akira
; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE
; TITLE OF INVENTION: LECTINS, DECTIN-1 AND DECTIN-2; COMPOSITIONS AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/772,440
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTXD:493
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-772-440-42

Alignment Scores:
Pred. No.: 2.56e-14 Length: 145
Score: 191.00 Matches: 30
Percent Similarity: 76.60% Conservatives: 6
Best Local Similarity: 63.83% Mismatches: 11
Query Match: 16.39% Indels: 0
DB: 3 Gaps: 0

US-09-766-511B-53 (1-209) x US-08-772-440-42 (1-145)
QY 163 PheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAlaSerIleValPheTrp 182
Db 2 TTCTGGCACCCCATGATGAACCCCAATCTCCAGAGAGGGGTGTTCATATAGTTTACTGG 61
QY 183 LysProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArgAsnSerIleCys 202
Db 62 AATCCTTCGAATGGGCTGGAATGATGTTTCTGTATAGTAAACACAATTCATATATG 121
QY 203 GluMetAsnLysIleTyrLeu 209
Db 122 GAAATGAAGAAGATTACCTA 142

Search completed: February 20, 2003, 20:40:40
```

GenCore version 5.1.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 20, 2003, 18:36:42 ; Search time 220 seconds  
(without alignments)  
2139.398 Million cell updates/sec

Title: US-09-766-511B-53  
Perfect score: 1165  
Sequence: 1 MMQEQQPOSTKRGWLSRL.....NDVICETRRNSICEMNKIYL 209

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Maximum Match 0%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_plus\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09766511/runat\_19022003\_094338\_29278/app\_query.fasta\_1.391  
-DB=N\_Geneseq\_101002 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTPMT=pto -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09766511@cgn\_1\_125\_@runat\_19022003\_094338\_29278 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1165	100.0	3114	22	AAS01375	Human TANGO 405 cd
2	892.5	76.6	850	20	AAZ07531	Human SDCMP3 polyp
3	793	68.1	630	20	AAZ07532	Mouse SDCMP3 polyp
4	793	68.1	1227	19	AAV42549	Mouse dectin-2 cdn
5	793	68.1	1252	22	AAS01378	Murine TANGO 405 a
6	740	63.5	821	22	AAS01376	Murine TANGO 405 c
7	679	58.3	501	19	AAV42554	Mouse dectin-2 ext
8	637	54.7	1312	22	AAF90241	Nucleotide sequenc
9	630	54.1	827	22	AAJ19729	Dendritic cell (DC
10	617	53.0	800	22	AAJ19730	Dendritic cell (DC
11	459.5	39.4	444	24	ABK52904	Human cdna encodin
12	459	39.4	402	24	ABK52903	Human cdna encodin
13	458	39.3	1091	22	AAS31385	Human cdna encodin
14	458	39.3	1091	24	ABQ66709	Human polynucleoti
15	458	39.3	1096	22	AAS31224	Human cdna encodin
16	458	39.3	1096	24	ABQ66548	Human polynucleoti
17	458	39.3	1104	20	AAK04865	Primate DCMPI C-le
c 18	458	39.3	1307	24	ABL90686	Human polynucleoti
c 19	439.5	37.7	582	24	ABK52904	Human dendritic ce
c 20	436	37.4	1418	20	AAK04867	rodent DCMPI C-lec
c 21	427.5	36.7	568	24	ABK52908	Human dendritic ce
c 22	419	36.0	558	24	ABK52907	Human dendritic ce
c 23	400.5	34.4	403	24	ABK52909	Human dendritic ce
c 24	398.5	34.2	968	22	AAF81743	Human membrane ass
c 25	398.5	34.2	997	20	AAK52274	Protein PRO244 cdn
c 26	398.5	34.2	997	22	AAK52936	Human DNA encoding
c 27	398.5	34.2	997	22	AAF72432	Human PRO244 cdna
c 28	390.5	33.5	2517	22	AAF66349	Murine cdna encodi
c 29	382	32.8	1013	24	ABK52905	Human dendritic ce
c 30	377.5	32.4	753	24	ABK52906	Human dendritic ce
c 31	375.5	32.4	561	24	ABK52911	Human dendritic ce
c 32	356.5	30.6	2059	21	AAA26386	Human secreted pro
c 33	356	30.6	324	22	AAK63162	Human immune/haema
c 34	303.5	26.1	323	22	AAK53734	Murine transport a
c 35	299.5	25.7	448	24	ABK52912	Human dendritic ce
c 36	290.5	24.9	666	24	ABK52910	Human dendritic ce
c 37	267.5	23.0	559	24	ABK52913	Human dendritic ce
c 38	254	21.8	1290	24	ABK63735	Rat sequence diffe
c 39	252.5	21.7	1309	24	ABN95737	Gene #2235 used to
c 40	252	21.6	1300	18	AAJ66951	Asialoglycoprotein
c 41	252	21.6	1300	24	ABN97236	Gene #3734 used to
c 42	251.5	21.6	1458	20	AAK04866	Primate DCMPI C-le
c 43	250	21.5	1370	20	AAK04868	Variant primate DC
c 44	250	21.5	1531	24	ABK94927	Human novel polynu
c 45	249	21.4	2005	24	ABL95574	Human angiogenesis

ALIGNMENTS

RESULT 1  
AAS01375  
ID AAS01375 standard; cdna: 3114 BP.  
XX  
AC AAS01375;  
XX  
DT 04-JUL-2001 (first entry)  
XX  
DE Human TANGO 405 cdna sequence.  
XX  
KW Human; TANGO 210; clone jthLai52H06; TANGO 364; TANGO 366; dectin-2;  
KW INTERCEPT 394; INTERCEPT 400; TANGO 405; cellular process regulator;  
KW gene therapy; growth modulator; proliferation; cell differentiation;  
KW lymphocyte; bone marrow cell migration; leukaemia; lymphoma;  
KW autoimmune disorder; ss.  
XX  
OS Homo sapiens.  
XX

FH Key Location/Qualifiers
FT CDS 154..783
FT /\*tag= a
FT /product= "TANGO 405 protein"
FT /note= "The ORF is specifically claimed"
FT sig\_peptide 154..297
FT mat\_peptide 298..780
FT /\*tag= b
FT /\*tag= c
XX
XX WO200118016-A1.
XX 15-MAR-2001.
XX
XX 30-JUN-2000; 2000WO-US18174.
XX
XX 10-SEP-1999; 99US-0393996.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Fraser CC, Sharp JD, Wrighton N, Myers PS, Goodearl ADJ;
XX
XX WPI; 2001-183280/18.
XX
XX P-PSDB; AAU00479.
XX
XX Isolated nucleic acid molecules encoding proteins useful as modulating
XX agents in regulating a variety of cellular processes are used for
XX treating e.g. cancer and autoimmune disorders -
XX
XX Claim 2; Fig 6A-6C; 326pp; English.
XX
XX The present sequence encoding for human TANGO 405 protein is isolated
XX from cDNA clone jthai152h06 from a human mixed lymphocyte reaction cDNA
XX library. It is 1 of 6 novel human proteins which include TANGO 210
XX (AAU00469), TANGO 364 (AAU00471), TANGO 366 (AAU00472), INTERCEPT 394
XX (AAU00473) and INTERCEPT 400 (AAU00476). Novel sequences for murine
XX TANGO 210 (AAU00470), INTERCEPT 400 (AAU00477), TANGO 405 (AAU00480) and
XX a rat INTERCEPT 400 (AAU00478) sequence are also described. The nucleic
XX acids encoding these novel proteins are useful as modulating agents in
XX regulating a variety of cellular processes and can be used to express
XX the proteins in a host cell in gene therapy applications. Human and
XX murine TANGO 405 proteins show sequence homology to murine dectin-2.
XX TANGO 405 modulates growth, proliferation, survival, differentiation,
XX activity, morphology and movement/migration of human lymphocytes and
XX bone marrow cells and tissues and can be used to prevent, diagnose or
XX treat leukaemia, lymphomas and autoimmune disorders.
XX
XX Sequence 3114 BP; 1001 A; 527 C; 517 G; 1069 T; 0 Other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 4e-116 Length: 3114
XX Score: 1165.00 Matches: 209
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 22 Gaps: 0
XX
XX US-09-766-511b-53 (1-209) x AAS01375 (1-3114)
XX
XX 1 MetMetGlnGlnGlnProGlnSerThrGluLysArgGlyTrpLeuSerLeuArgLeu 20
XX
XX 154 ATGATGCAAGAGCAGCAACCTCAAGTACAGAGAAAAGAGGCTGGTGTCCCTGAGACTC 213
XX
XX 21 TrpSerValAlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysVal 40
XX
XX 214 TGGTCTGTGGCTGGGATTCCATTGCACCTCTCAGTCTGCTTCATTGTGAGCTGTGTA 273
XX
XX 41 ValThrTyrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHisSerTyr 60
XX
XX 274 GTAACCTACCATTTTACATATGGTGAACCTGGCAAAAGGCTGTCTGAACCTACATCAT 333
XX
XX 61 HisSerSerLeuThrCysPheSerGluGlyThrLysValProAlaTrpGlyCysCysPro 80
XX
XX

Db 334 CATTCAAGTCTCACCTTCACTTCACTGAGGACAAAGGTGCCAGCCTGGGGATGTTGCCCA 393
QY 81 AlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGluLysValTrp 100
Db 394 GCITCTTGGGAAGTCATTGGTTCCAGTTGCTACTTCAATTCAGTGAAGAGAAGGTTTGG 453
QY 101 SerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsnThrGlu 120
Db 454 TCTAAGAGTGCAGCAGAACTCTGTGAGATGGGAGCACATTTGGTTGTTCACACAGAA 513
QY 121 AlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeu 140
Db 514 GCAGACAGAAATTCATTGTCCAGCAGCTGAATGAGTCAATTTCTTATTTCTGGGGCTT 573
QY 141 SerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsn 160
Db 574 TCAGACCCACAAAGTAAATAATTGGCAATGGATTGATAGACACACTTATGAGAAAAAT 633
QY 161 ValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAlaSerIleVal 180
Db 634 GTCAGATTTTGGCACCCTAGGTGAGCCCAATCAATTCGCAGACGAATGTCTTCAATAGTC 693
QY 181 PheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArgAsnSer 200
Db 694 TTTCTGAAACCTACAGATGGGCTGGAATGATGATTTATCTGTGAAACTAGAAAGGAATTC 753
QY 201 IleCysGluMetAsnLysIleTyrLeu 209
Db 754 ATATGTGAGATGAATAAGATTTACCTA 780
XX
XX RESULT 2
XX AAZ07531
XX ID AAZ07531 standard; cDNA; 850 BP.
XX AC AAZ07531;
XX XT 26-NOV-1999 (first entry)
XX DE Human SDCMP3 polypeptide encoding cDNA.
XX KW Binding compound; antibody binding site; primate; rodent; SDCMP3; SDCMP4;
XX KW Schering dendritic cell membrane protein; dendritic cell physiology;
XX KW genetic fingerprinting; cancer immunotherapy; abnormal proliferation;
XX KW cancer; forensics; human; lectin 73; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT CDS 108..596
XX FT /\*tag= a
XX FT /product= "SDCMP3"
XX
XX WO9947673-A2.
XX
XX 23-SEP-1999.
XX
XX 16-MAR-1999; 99WO-US03740.
XX
XX 17-MAR-1998; 98US-0040111.
XX
XX (SCHE ) SCHERING CORP.
XX
XX Chalus L, Quan AB, Bates EEM, Gorman DM, Saeland S, Lebecque SJE;
XX PI Philipps JH;
XX
XX WPI; 1999-562114/47.
XX P-PSDB; AAY27446.
XX
XX Binding compound specific for primate or rodent Schering dendritic cell
XX membrane proteins -
XX
XX Claim 8; Page 80; 89pp; English.
XX

CC The invention relates to a binding compound comprising an antibody  
CC binding site which specifically binds to: (a) a primate or rodent  
CC Schering dendritic cell membrane protein 3 (SDCMP3) protein; or (b) a  
CC primate SDCMP4 protein. The binding compound can be used to modulate  
CC dendritic cell physiology or function. The sequences may also be used as  
CC probes in forensic techniques, such as genetic fingerprinting. They can  
CC also be used to distinguish tissue and cell types in situ or in vitro.  
CC The sequences may also be useful in cancer immunotherapy and in the  
CC treatment of conditions associated with abnormal physiology or  
CC development, including abnormal proliferation, e.g. cancerous conditions,  
CC or degenerative conditions. The present sequence represents a cDNA  
CC encoding the human SDCMP3 (lectin 73).  
XX

SQ Sequence 850 BP; 257 A; 153 C; 188 G; 252 T; 0 other;

## Alignment Scores:

Pred. No.: 2,54e-87 Length: 850  
Score: 892.50 Matches: 165  
Percent Similarity: 97.63% Conservative: 0  
Best Local Similarity: 97.63% Mismatches: 1  
Query Match: 76.61% Indels: 3  
DB: 20 Gaps: 1

US-09-766-511B-53 (1-209) x AAZ07531 (1-850)

QY 1 MetMetGlnGluGlnProGlnSerThrGluLysArgGlyTrpLeuSerLeuArgLeu 20  
DB 108 ATGATGCAAGCAGCAGCAACCTCAAGTACAGAGAAAGAGCGTGTTCCTCCAGACTC 167  
QY 21 TrpSerValAlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysVal 40  
DB 168 TGGTCTGTGGCTGGGATTTCCATTGCACCTCTCAGTGTCTTCATTGTGAGCTGTGTA 227  
QY 41 ValThrTrpHisPheThrTrpGlyGluThrGlyLysArgLeuSerGluLeuHisSerTyr 60  
DB 228 GTAACTTACCAATTTACATATGTTGAACCTGGCAAAAGGCTGTCTGAACCTACACTCATAT 287  
QY 61 HisSerSerLeuThrCysPheSerGluGlyThrLysValProAlaTrpGlyCysCysPro 80  
DB 288 CATTCAAGCTTACCTGCTTCAAGTGAAGGACAAAGTGCCAGCTGGGGATGTCCCA 347  
QY 81 AlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGluGluLysValTrp 100  
DB 348 GCTTTCTTGGGAAGTCATTTGGTTCCAGTGTCTACTTTCATTCCAGTGAAGAGAGTTTG 407  
QY 101 SerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsnThrGlu 120  
DB 408 TCTAAGAGTCAGCAGAACTGTGTGAGATGGGAGCACATTTGGTTGTTCACACAGAA 467  
QY 121 AlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTrpPheLeuGlyLeu 140  
DB 468 GCAGAGCAGAAATTTCAATGTCACAGCAGCTGAATGAGTCATTTTCTTATTTCTGGGCTT 527  
QY 141 SerAspProGlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrProTrpGluLysAsn 160  
DB 528 TCAGACCCACAGGTAATAATAATTTGGCAATGGATTTGATAAGACACCTTATGAGAAAAAT 587  
QY 161 ValArg-----PheTrpHisLeu 166  
DB 588 GTCAGGTGAGTCAGTCTCTGGGCGCTT 614

## RESULT 3

AAZ07532

ID AAZ07532 standard; cDNA; 630 BP.

XX

AC AAZ07532;

XX

DT 26-NOV-1999 (first entry)

XX

DE Mouse SDCMP3 polypeptide encoding cDNA.

XX

KW Binding compound; antibody binding site; primate; rodent; SDCMP3; SDCMP4;

KW Schering dendritic cell membrane protein; dendritic cell physiology;

KW genetic fingerprinting; cancer immunotherapy; abnormal proliferation;  
KW cancer; forensic; mouse; ss.

XX Mus sp.

XX Key Location/Qualifiers  
FH 1..630  
FT CDS /\*tag= a  
FT /product= "SDCMP3"

XX WO9947673-A2.

XX 23-SEP-1999.

XX 16-MAR-1999; 99WO-US03740.

XX 17-MAR-1998; 98US-0040111.

XX (SCHE ) SCHERING CORP.

PI Chalus L, Quan AB, Bates EEM, Gorman DM, Saeland S, Lebecque SJE;  
PI Philipps JH;

DR WPI; 1999-562114/47.  
DR P-PSDB; AAY27447.

PT Binding compound specific for primate or rodent Schering dendritic cell  
PT membrane proteins

XX Claim 8; Page 82; 89pp; English.

CC The invention relates to a binding compound comprising an antibody  
CC binding site which specifically binds to: (a) a primate or rodent  
CC Schering dendritic cell membrane protein 3 (SDCMP3) protein; or (b) a  
CC primate SDCMP4 protein. The binding compound can be used to modulate  
CC dendritic cell physiology or function. The sequences may also be used as  
CC probes in forensic techniques, such as genetic fingerprinting. They can  
CC also be used to distinguish tissue and cell types in situ or in vitro.  
CC The sequences may also be useful in cancer immunotherapy and in the  
CC treatment of conditions associated with abnormal physiology or  
CC development, including abnormal proliferation, e.g. cancerous conditions,  
CC or degenerative conditions. The present sequence represents a cDNA  
CC encoding the mouse SDCMP3.

SQ Sequence 630 BP; 171 A; 135 C; 152 G; 172 T; 0 other;

## Alignment Scores:

Pred. No.: 1.07e-76 Length: 630  
Score: 793.00 Matches: 145  
Percent Similarity: 77.00% Conservative: 19  
Best Local Similarity: 68.08% Mismatches: 41  
Query Match: 68.07% Indels: 8  
DB: 20 Gaps: 4

US-09-766-511B-53 (1-209) x AAZ07532 (1-630)

QY 1 MetMetGlnGluGlnProGlnSerThrGluLysArgGly-----TrpLeuSerLeu 18  
DB 1 ATGGTGCAGGAAGACAATCCAGGG-----AAGGAGCTGTCTGG---ACCCTG 48  
QY 19 ArgLeuTrpSerValAlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSer 38  
DB 49 AGACTCTGTCAGCTGCTGTGATTTCCATTTCTACTCTTTGAGTACCTGTTTCTTCGAGC 108  
QY 39 CysValValThrTrpHisPheThrTrpGlyGluThrGlyLysArgLeuSerGluLeuHis 58  
DB 109 TGTGTGGTGACTTACCAATTTATTATGACCCAGCCAGCTAGAGACTATATCACTTCC 168  
QY 59 SerTrpHisSerSerLeuThrCysPheSerGluGlyThrLysVal-----ProAlaTrp 76  
DB 169 ACATACCAATTTCCAGTCTCACCTGCTTCAGTGAAGGAGACTATGGTCTCAGAAAAATGTGG 228  
QY 77 GlyCysCysProAlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGlu 96

Db 229 GGATGCTGCCAAATCACTGGAAGTCATTTGGCTCCAGCTGCTACCTCATTTCTACCAAG 288  
 QY 97 GluLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValVal 116  
 Db 289 GAGAACTTCTGGAGCACCAGTCAGCAGAACCTGTGTTGATGGGGCTCATCTGGTGGTG 348  
 QY 117 PheAsnThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyr 136  
 Db 349 ATCAATACTGAAGCGGAGCAGAAATTTTCATACCCAGCAGCTGAATGAGTCACCTTCTTAC 408  
 QY 137 PheLeuGlyLeuSerAspProGlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrPro 156  
 Db 409 TTCTTGGGTCTTTCGGATCCCAAGGTAATGGCAATGGCAATGGATGATGATCTCT 468  
 QY 157 TyrGluLysAsnValArgPheThrPheHisLeuGlyGluProAsnHisSerAlaGluGlnCys 176  
 Db 469 TTCAGTCAAAATGTCAGGTTCTGGCACCCCATGAACCAATCTTCAGAGAGCGGTGT 528  
 QY 177 AlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThr 196  
 Db 529 GTTTCATAGTTTACTGGAATCCTTCGAAATGGGGCTGGAATGATGTTTCTGTGATAGT 588  
 QY 197 ArgArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209  
 Db 589 AAACACAATTCATATGTGAAATGAAGAAGATTACCTA 627  
 RESULT 4  
 AAV42549  
 ID AAV42549 standard; cDNA; 1227 BP.  
 AC AAV42549;  
 DT 09-NOV-1998 (first entry)  
 DE Mouse dectin-2 cDNA.  
 KW Dectin-2; dendritic cell; c-type lectin; mouse; immunity; adjuvant;  
 KW allergy; autoimmune disease; gene therapy; vaccine; diagnosis;  
 KW drug screening; ss.  
 OS Mus sp.  
 FH Key  
 FT CDS  
 FT Location/Qualifiers  
 FT 146..775  
 FT /\*tag= a  
 PN WO9828332-A2.  
 PD 02-JUL-1998.  
 PF 22-DEC-1997; 97WO-US23761.  
 PR 20-DEC-1996; 96US-0772440.  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 PI Ariizumi K, Takashima A;  
 PI WPI: 1998-377594/32.  
 DR P-PSDB; AAW63010.  
 XX Nucleic acid encoding dendritic cell specific peptide(s) dectin-1  
 PT and -2 - useful, e.g. to regulate immune response, as vaccine  
 PT adjuvants, for diagnosis and drug screening  
 XX Claim 43; Page 141-142; 200pp; English.  
 XX This cDNA clone codes for a novel murine dendritic cell (DC)  
 CC specific member of the C-type lectin family, termed DC-associated  
 CC C-type lectin-2, or dectin-2 (see AAW63010), that is essential for  
 CC DC-mediated T cell activation. Dectin-1 cDNA (see AAV42548) has  
 CC also been identified. The invention provides: dectin-1 and

CC dectin-2 polypeptides (see also AAW63009-22 and AAW69236-37), useful  
 CC for purifying T cells, for detecting autoantibodies and for  
 CC up-regulating immunity e.g. as vaccine adjuvants; dectin DNA (see  
 CC AAV42548-54, AAV42558-60 and AAV44850-51); expression vectors;  
 CC recombinant host cells; probes and primers useful e.g. for identifying  
 CC human dectin-like molecules (see AAV42560); antibodies; compounds that  
 CC modulate dectin-mediated activation of T cells; transgenic animals  
 CC useful for studying dectin function and for drug screening; and  
 CC dectin ligands. Dectin expression can be downregulated by  
 CC antisense sequences or ribozymes, or by inactivating the dectin  
 CC genes by homologous recombination, e.g. abrogation of dectin  
 CC expression is useful for treating allergy and autoimmune  
 CC disease. Alternatively gene disruption is done in vitro, the  
 CC cells pulsed with an antigen, and returned to the patient to  
 CC provide long-term non-responsiveness to the antigen. Since DC are  
 CC 'professional' antigen-presenting cells, genetic vaccines targeted  
 CC to them should not induce tolerance.  
 XX Sequence 1227 BP; 335 A; 271 C; 276 G; 345 T; 0 other;  
 SQ

Alignment Scores:  
 Pred. No.: 2,82e-76 Length: 1227  
 Score: 793.00 Matches: 145  
 Percent Similarity: 77.00% Conservative: 19  
 Best Local Similarity: 68.08% Mismatches: 41  
 Query Match: 68.07% Indels: 8  
 DB: 19 Gaps: 4

US-09-766-511b-53 (1-209) x AAV42549 (1-1227)  
 QY 1 MetMetGlnGlnGlnProGlnSerThrGluLysArgGly-----TrpLeuSerLeu 18  
 Db 146 ATGGTGCAGGAAGACAAATCCCAAGG-----AAGGGAGCTGCTGG---ACCCTG 193  
 QY 19 ArgLeuTrpSerValAlaGlyIleSerIleAlaLeuSerAlaCysPheIleValSer 38  
 Db 194 AGACTGTGGTCAGCTGCTGTGATTTCCATGTTACTTCTGAGTACCTGTTTCATGGAGC 253  
 QY 39 CysValValThrTyrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHis 58  
 Db 254 TGTGTGTGACTTACCAATTTATTTATGGACCCAGTGAAGACTATATGAACCTAC 313  
 QY 59 SerTyrHisSerSerLeuThrCysPheSerGluGlyThrLysVal-----ProAlaTrp 76  
 Db 314 ACATACCATTTCCAGTCTCACCTGCTTCAGTGAAGGAGCTATGCTGCAGAAAAATGTGG 373  
 QY 77 GlyCysCysProAlaSerTrpLysSerPheGlySerCysTyrPheIleSerSerGlu 96  
 Db 374 GGATGCTGCCCAATCACTGGAAGTCATTTGGCTCCAGCTGCTACCTCATTTACCAAG 433  
 QY 97 GluLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValVal 116  
 Db 434 GAGAACTTCTGGAGCACCAGTCAGCAGAACCTGTGTCAGATGGGGCTCATCTGGTGGTG 493  
 QY 117 PheAsnThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyr 136  
 Db 494 ATCAATACTGAAGCGGAGCAGAAATTTTCATCACCAGCAGCTGAATGAGTCACCTTCTTAC 553  
 QY 137 PheLeuGlyLeuSerAspProGlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrPro 156  
 Db 554 TTCTTGGGTCTTTCGGATCCCAAGGTAATGGCAATGGCAATGGATGATGATCTCT 613  
 QY 157 TyrGluLysAsnValArgPheThrPheHisLeuGlyGluProAsnHisSerAlaGluGlnCys 176  
 Db 614 TTCAGTCAAAATGTCAGGTTCTGGCACCCCATGAACCAATCTTCAGAGAGCGGTGT 673  
 QY 177 AlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThr 196  
 Db 674 GTTTCATAGTTTACTGGAATCCTTCGAAATGGGGCTGGAATGATGTTTCTGTGATAGT 733  
 QY 197 ArgArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209  
 Db 734 AAACACAATTCATATGTGAAATGAAGAAGATTACCTA 772





XX 15-MAR-2001.  
 PD  
 XX  
 XX 30-JUN-2000; 2000WO-US18174.  
 PF  
 XX  
 XX 10-SEP-1999; 99US-0393996.  
 PR  
 XX  
 XX (MILL-) MILLENNIUM PHARM INC.  
 PA  
 XX  
 XX Fraser CC, Sharp JD, Wrighton N, Myers PS, Goodearl ADJ;  
 PI  
 XX WPI; 2001-183280/18.  
 XX P-PSDB; AAU00480.  
 DR  
 DR  
 XX Isolated nucleic acid molecules encoding proteins useful as modulating  
 PT agents in regulating a variety of cellular processes are used for  
 PT treating e.g. cancer and autoimmune disorders -  
 PT  
 XX Disclosure; Fig 6E-6F; 326pp; English.  
 PS  
 XX  
 CC The present sequence encoding for murine TANGO 405 protein is isolated  
 CC from cDNA clone jtmMa025a11 from a long-term bone marrow cDNA library.  
 CC An alternative splice variant of this protein is also described  
 CC (AAU00482). Other novel murine proteins include TANGO 210 (AAU00470) and  
 CC INTERCEPT 400 (AAU00477). Six novel human proteins which include  
 CC TANGO 210 (AAU00469), TANGO 364 (AAU00471), TANGO 366 (AAU00472),  
 CC INTERCEPT 394 (AAU00473), INTERCEPT 400 (AAU00476) and TANGO 405  
 CC (AAU00479), and a rat INTERCEPT 400 (AAU00478) sequence are also  
 CC described. The nucleic acids encoding these novel proteins are useful as  
 CC modulating agents in regulating a variety of cellular processes and can  
 CC be used to express the proteins in a host cell in gene therapy  
 CC applications. Human and murine TANGO 405 proteins show sequence homology  
 CC to murine lectin-2. TANGO 405 modulates growth, proliferation, survival,  
 CC differentiation, activity, morphology and movement/migration of human  
 CC lymphocytes and bone marrow cells and tissues and can be used to prevent,  
 CC diagnose or treat leukaemia, lymphomas and autoimmune disorders.  
 CC  
 XX Sequence 821 BP; 214 A; 188 C; 195 G; 223 T; 1 other;  
 SQ

Alignment Scores:  
 Pred. No.: 9,09e-71 Length: 821  
 Score: 740.00 Matches: 142  
 Percent Similarity: 75.94% Conservative: 19  
 Best Local Similarity: 66.98% Mismatches: 42  
 Query Match: 63.52% Indels: 10  
 DB: 22 Gaps: 4

US-09-766-511B-53 (1-209) x AAS01376 (1-821)  
 QY 1 MetMetGlnGlnGlnProGlnSerThrGluLysArgGly-----TrpLeuSerLeu 18  
 DB 174 ATGGTGCAGGAAGACAATCCCAAGG-----AAGGGAGTCTGCTGG---ACCCGTG 221  
 QY 19 ArgLeuTrpSerValAlaGlyLeuSerIleAlaLeuLeuSerAlaCysPheIleValSer 38  
 DB 222 AGACTCTGTGCTGCTGTGATTTCCAGTTACTCTTGAGTACCTGTTTCATTGCGAGC 281  
 QY 39 CysValValThrTyrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHis 58  
 DB 282 TCTGTGGTGACTTACCAATTTATTATGACCCAGCCAGCCATAGAGACTATATGAATTTCAC 341  
 QY 59 SerTyrHisSerSerLeuThrCysPheSerGluGlyThrLysVal-----ProAlaTrp 76  
 DB 342 ACATACCACTTCAGTCTCACCTCTTCAGTGAAGGGACTATGGTGTGAGAAAAATGTGG 401  
 QY 77 GlyCysCysProAlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGlu 96  
 DB 402 GGATGCTGCCCAAACTACCTGGAGTCAITTTGGCTCCAGCTACTCTCATTTTACCAAG 461  
 QY 97 GluLysValTrpSerLysSerGluGlnAsnGlnValGluMetGlyAlaHisValVal 116  
 DB 462 GAGAACTTCTGGAGCACCAGTCAGCAGAACTCTGTTTCAGATGGGGCTCTCTGTGGTG 521

QY 117 PheAsnThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyr 136  
 DB 522 ATCAATACTGAAGCGGAGCAGAAATTTCAATCCAGCAGCTGAATGAGTCACTTCTTAC 581  
 QY 137 PheLeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrPro 156  
 DB 582 TTCCTGGGTCTTCGGATCCCAA-GGTAATGGCAATGGCAATGGATCGATGATCTCCT 640  
 QY 157 TyrGluLysAsnValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCys 176  
 DB 641 TTCAGTCAAAATGTCAGGTTCTGGCACCCCATCAACCCCAATCTTCCAGAGAGCGGTGT 700  
 QY 177 AlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluTh 196  
 DB 701 GTTTCATAGTTTACTGGAATCCTTCGAAATGGGCTGGGAATCATCTTTCTGTGATAG 760  
 QY 196 rArgArgAsnSerIleCysGluMetAsnLysIle 207  
 DB 761 TAAACACAATTCATATGTGAAATGAANAAGATT 794  
 RESULT 7  
 AAV42554  
 ID AAV42554 standard; cDNA; 501 BP.  
 XX  
 AC AAV42554;  
 XX  
 DT 09-NOV-1998 (first entry)  
 XX  
 DE Mouse lectin-2 extracellular domain cDNA.  
 XX  
 KW Dectin-2; dendritic cell; c-type lectin; mouse; immunity; adjuvant;  
 KW allergy; autoimmune disease; gene therapy; vaccine; diagnosis;  
 KW drug screening; ss.  
 XX  
 OS Mus sp.  
 XX  
 PN W09828332-A2.  
 XX  
 PD 02-JUL-1998.  
 XX  
 PF 22-DEC-1997; 97WO-0523761.  
 XX  
 PR 20-DEC-1996; 96US-0772440.  
 XX  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 XX  
 PI Arizumi K, Takashima A;  
 XX  
 DR WPI; 1998-377594/32.  
 DR P-PSDB; AAW63022.  
 XX  
 PT Nucleic acid encoding dendritic cell specific peptide(s) dectin-1  
 PT and -2 - useful, e.g. to regulate immune response, as vaccine  
 PT adjuvants, for diagnosis and drug screening  
 PS  
 PS Disclosure; Page 153; 200pp; English.  
 XX  
 CC This cDNA sequence codes for the extracellular domain (see  
 CC AAW63022) of a novel murine dendritic cell (DC) specific member of  
 CC the C-type lectin family, termed DC-associated C-type lectin-2, or  
 CC dectin-2 (see also AAW63010), that is essential for DC-mediated T  
 CC cell activation. Full-length cDNA (see AAV42549) was isolated by  
 CC subtractive hybridisation from BALB/c mouse epidermis DC line XS52.  
 CC The invention provides: dectin-1 and -2 polypeptides (see also  
 CC AAW63009-22 and AAW69236-37), useful for purifying T cells, for  
 CC detecting autoantibodies and for up-regulating immunity e.g. as  
 CC vaccine adjuvants; dectin DNA (see AAV42548-54, AAV42558-60 and  
 CC AAV44850-51); expression vectors; recombinant host cells; probes  
 CC and primers; antibodies; compounds that modulate dectin-mediated  
 CC activation of T cells; transgenic animals; and dectin ligands.  
 CC Dectin expression can also be down-regulated to treat allergy and  
 CC autoimmune disease. Recombinant extracellular domain has been  
 CC expressed as a His-tagged protein (see AAW69237) useful for antibody

```

CC production.
XX Sequence 501 BP; 142 A; 108 C; 115 G; 136 T; 0 other;
SQ
Alignment Scores:
pred. No.: 1.9e-64 Length: 501
score: 679.00 Matches: 118
Percent Similarity: 79.04% Conservative: 14
Best Local Similarity: 70.66% Mismatches: 33
Query Match: 58.28% Indels: 2
DB: 19 Gaps: 1

```

US-09-766-511B-53 (1-209) x AAV42554 (1-501)

Qy	45	PheThrTyrglyGluThrGlyLysArgLeuSerGluLeuHisSerTyrHisSerSerLeu	64
Db	1	TTTATTATGGACAGCCAGCAGTAGAAGACTATATGAATCTTCACATACCATTCAGTCTC	60
Qy	65	ThrCysPheSerGluGlyThrLysVal-----ProAlaIrrpGlyCysCysProAlaSer	82
Db	61	ACCTGCTTCAGTGAAGGACTATGGTGTCTCAGAAAAATGTGGGATGCTGCCAAATCAC	120
Qy	83	TrpLysSerPheGlySerSerCysTyrPheIleSerSerGluGluLysValTrpSerLys	102
Db	121	TGGAAGTCATTGGCTGCAGCTGCTACTCATTTCTACCAAGGAGAACTTCTGGAGACC	180
Qy	103	SerGluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsnThrGluAlaGlu	122
Db	181	AGTGACGACACTGTTTCAGATGGGGCTCATCTGGTGGTATCAATCACTGAAGCGAG	240
Qy	123	GlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeuSerAsp	142
Db	241	CAGAAATTTTCATCACCACGACGCTGAATGAGTCACCTTTCTTACTCTGGGCTTTTCGGAT	300
Qy	143	ProGlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsnValArg	162
Db	301	CCACAGGTATGGCAATGGCAATGGATCGATGATACCTCTTCAGTCAAAATGTCAGG	360
Qy	163	PheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAlaSerIleValPheTrp	182
Db	361	TTCTGGCACCCCATGAACCCCAATCTTCCAGAGAGCGGTGTGTTCATTAATAGTTACTGG	420
Qy	183	LysProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArgAsnSerIleCys	202
Db	421	AATCCTTCGAATGGGGCTGGAATGTTTCTGTGATAGTAAACAAATTCATATGT	480
Qy	203	GluMetAsnLysIleTyrLeu	209
Db	481	GAATGAAGAAGATTTACCTA	501

RESULT 8  
AAAF90241  
ID AAAF90241 standard; cDNA: 1312 BP.

AC AAF90241:

XX  
DT 06-AUG-2007 (first entry)

XX Nivaloxido sequence of BDC

[illegible]

KW viral infection; autoimmune disease; allergic response; cancer; ss. ....

OS Homo sapiens.

FH	Key	Location/Qualifiers
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

```
FT      /*tag= a
```

XX  
XX

XX  
XX  
1040CT00ZOM N3  
12Z / 1040CT00ZOM N3

FD 207-141-23 MAY 23 1967

XX PD

15-NOV-2000; 2000WO-IB01832.

15-NOV-1999: 99US-0165555.

23-NOV-1999: 990US-0167076;  
28-JAN-2000: 2000US-0179003;

07-FEB-2000; 2000US-0180775.  
11-APR-2000; 2000US-0196824

13-APR-2000; 2000US-0197205.

(MILT-) MILTENYI BIOTECH GMBH

Schmitz J, Dzionek A, Buck

WPI; 2001-355622/37.

use of antigen-binding frag

allergic responses and cancer

Claim 134; Fig 12; 115pp; En

The present sequence encodes

enriched in dendritic cells

antigen-binding fragments specifically obtained are used to treat various

response, and cancer. BDCA-1 antibodies and their antigen

enumerate and isolate DC pop-

COCHINGO 1212 RD. 274 A. 301

Alignment Scores:

Pred. No.:	2.81e-59	Length:	1312
Score:	637.00	Matches:	115
Percent Similarity:	68.08	Conservative:	30
Best Local Similarity:	53.99%	Mismatches:	64
Query Match:	54.68%	Indels:	54
DB:	22	Gaps:	2

US-09-766-511B-53 (1-209) x AAF90241 (1-1312)

QY 1 MetMetClnGluGlnGlnProGlnSerThrClnLys---ArgGlyTrpLeuSerLeuArg 19

491 ATGGTGCCCTGAAGAGAGAGAGCCCTCAAGACCCGAGAGAGAAAGGACCTCTGGTGGTCTCCAGATTTGAAG

Db 551 GTCTGGTCCATGGCACTCGTATCCATCTTGCTCCTCAGTGTCTGTTTCACTGTGAGTTCT 610

[illegible]

Qy 58 ---HisSerTyrHisSerSerLeuThrCysPheSerGluGlyThrLysValProAlaTrp 76  
:::

QY	77	GlyCysCysProAlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGlu	96
----	----	--	----

Db 731 AGTGTGCCCAACCCCTTGGACTTCATTTCACTAGTTGCTACTTTATTTCTACTGGG 790

Db 791 ATGCAATCTTGGACAAGAGTCAAAAGAACTGTTCTGTGATGGGGGCTGATCTGGTGGTG 850

Qy 117 PheAsnThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyr 136

.....

QY 137 PheLeuGlyLeuSerAspProGlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrPro 156  
 DB 911 TTTCTGGGCTGTACAGTCCAGGGGTGGCGACATTTGGCAATGGTGGACGACACCA 970  
 QY 157 TyrGluLysAsnValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCys 176  
 DB 971 TACAATGAAATGTACATCTCGGCACCTCAGGTGAACCAATAACCTTGATGAGCGTTGT 1030  
 QY 177 AlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThr 196  
 DB 1031 GCGATAATAAATTCGTTCTTCAGAGAAGTGGGGCTGGAATGACATTCACCTGTCATGTA 1090  
 QY 197 ArgArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209  
 DB 1091 CCTCAGAGTCAATTTGCAAGATGAAGAAGATCTACATA 1129  
 RESULT 9  
 AAD19729  
 ID AAD19729 standard; DNA; 827 BP.  
 XX  
 AC AAD19729;  
 XX  
 DT 18-DEC-2001 (first entry)  
 XX  
 DE Dendritic cell (DC) DCLEC gene.  
 XX  
 KW Dendritic cell; DC; DCLEC protein; gene therapy; dermatological; vaccine;  
 KW atopic dermatitis; autoimmune disease; inflammatory skin disease; cancer;  
 KW immunosuppressive; AIDS; Acquired immune deficiency syndrome; cytostatic;  
 KW chromosomal identification; pharmaceutical; hypersensitivity; virucide;  
 KW transplant rejection; chronic inflammatory disease; anti-HIV; ds.  
 XX  
 OS Unidentified.  
 XX  
 FH Key  
 FT Location/Qualifiers  
 FT CDS  
 FT 1..657  
 FT /\*tag= a  
 FT /product= "Dendritic cell DCLEC protein"  
 XX  
 PN WO200172773-A2.  
 XX  
 PD 04-OCT-2001.  
 XX  
 PF 28-MAR-2001; 2001WO-EP03542.  
 XX  
 PR 29-MAR-2000; 2000US-192934P.  
 PR 18-MAY-2000; 2000US-205020P.  
 PR 18-MAY-2000; 2000US-205026P.  
 PR 19-MAY-2000; 2000US-205767P.  
 PR 19-MAY-2000; 2000US-205769P.  
 XX  
 PA (NOVS ) NOVARTIS AG.  
 PA (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.  
 XX  
 PI Werner G, Phares W, Jaritz M, Lapp H, Kalthoff rs;  
 XX  
 DR WPI; 2001-616466/71.  
 DR P-PSDB; AAE12079.  
 XX  
 DR New polypeptides for screening therapeutic agonists and antagonists  
 PT comprise dendritic cell polypeptides -  
 XX  
 PS Claim 1; Page 50-51; 52pp; English.  
 XX  
 CC The invention relates to dendritic cell (DC) proteins and their  
 CC corresponding DNA molecules. A pharmaceutical composition comprising  
 CC agonist and antagonist of DC proteins are useful for treating abnormal  
 CC conditions related to both an excess of and insufficient level of  
 CC expression of DC gene, or related to both an excess of and insufficient  
 CC activity of DC protein. Soluble form of DC proteins are used as an active  
 CC ingredient in combination with pharmaceutical acceptable carriers.  
 CC DC genes and proteins are useful for treating chronic inflammatory

CC diseases, autoimmune diseases, transplant rejection crisis, including  
 CC inflammatory skin diseases such as contact hypersensitivity, atopic  
 CC dermatitis or virally-induced immune suppression such as AIDS and cancer.  
 CC DC protein is useful for inducing immunological response in a mammal, and  
 CC as immunogen to produce antibodies immunospecific for the polypeptide.  
 CC DC gene is useful in gene therapy. DC gene is also useful as a diagnostic  
 CC reagent, and for chromosomal identification. The present sequence is  
 CC dendritic cell (DC) DCLEC gene which is found to belong to the family  
 CC of C-type lectins with one single carbohydrate recognition domain at the  
 CC C-terminal end.  
 XX

SQ Sequence 827 BP; 242 A; 164 C; 188 G; 233 T; 0 other;

Alignment Scores:  
 Pred. No.: 8 31e-59 Length: 827  
 Score: 630.00 Matches: 114  
 Percent Similarity: 67.94% Conservatives: 28  
 Best Local Similarity: 54.55% Mismatches: 63  
 Query Match: 54.08% Indels: 4  
 DB: 22 Gaps: 2

US-09-766-511B-53 (1-209) x AAD19729 (1-827)

QY 5 GlnGlnProGlnSerThrGluLys---ArgGlyTrpLeuSerLeuArgLeuTrpSerVal 23  
 DB 28 CAAGAGCCTCAGACCGGAGAGAAAGGACTCTGGTGGTTCAGTTGAAGCTCTGGTCCATG 87  
 QY 24 AlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysValValThrTyr 43  
 DB 88 GCAGTCGTATCCATCTGTCTCCCTCAGTGTCTGTTTCTCAGTGTGTCTGGTGGCTCAC 147  
 QY 44 HisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeu-----HisSerTyr 60  
 DB 148 AATTATTATGATAGCAAAACCTGTCAAGAGGCTGTCCAAAGTTACGACAGATCAACAGTAT 207  
 QY 61 HisSerSerLeuThrCysPheSerGluGlyThrLysValProAlaTrpGlyCysCysPro 80  
 DB 208 CATCAAGCCTGACCTGCCTCATGGAAGGAAAGGACATAGAGATTGGAGCTGCTGCCCA 267  
 QY 81 AlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGluGluLysValTrp 100  
 DB 268 ACCCTTGGACTTCATTTAGTCTAGTTGCTACTTTATTCTACTGGGATGCAATCTTGG 327  
 QY 101 SerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsnThrGlu 120  
 DB 328 ACTAAGAGTCAAAAGAACTGTTCTGTGTATGGGGGCTGATCTGTGTGTATCAACACAGG 387  
 QY 121 AlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeu 140  
 DB 388 GAAGAACAGGATTTTCATCATCTCAGAATCTGAAAGAAATTTCTTTATTTCTGGGGCTG 447  
 QY 141 SerAspProGlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsn 160  
 DB 448 TCAGATCCAGGGGTGGCGACATTTGGCAATGGTGTGACACAGACACATACATATAAAT 507  
 QY 161 ValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAlaSerIleVal 180  
 DB 508 GTCACATTTCTGGCACCACAGTGAACCAATAACCTTGTATGAGCGTTGTGCGATAATAAT 567  
 QY 181 PheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArgAsnSer 200  
 DB 568 TTCCTGTTCTTCAGAGAAGTGGGGCTGGAATGACATTCACCTGTCTACCTCAGAGTCA 627  
 QY 201 IleCysGluMetAsnLysIleTyrLeu 209  
 DB 628 ATTTGCAAGATGAAGAAGATCTACATA 654  
 RESULT 10  
 AAD19730  
 ID AAD19730 standard; DNA; 800 BP.  
 XX  
 AC AAD19730;  
 XX

18-DEC-2001 (first entry)  
Dendritic cell (DC) DCLEC/SPLICE 1 variant gene.  
Dendritic cell; DC; DCLEC protein; gene therapy; dermatological; vaccine; atopic dermatitis; autoimmune disease; inflammatory skin disease; cancer; immunosuppressive; AIDS; Acquired immune deficiency syndrome; cytostatic; chromosomal identification; pharmaceutical; hypersensitivity; virucide; transplant rejection; chronic inflammatory disease; anti-HIV; variant; ds.  
Unidentified.  
Key CDS Location/Qualifiers  
1..634  
/\*tag= a  
/product= "Dendritic cell DCLEC/SPLICE 1 variant protein"  
/trans\_except= (pos:1, aa:Pro)  
/note= "This codon has an apparent deletion of 2 nucleotides, which alters the reading frame; CDS does not include start codon"  
/partial  
W0200172773-A2.  
04-OCT-2001.  
28-MAR-2001; 2001WO-EP03542.  
29-MAR-2000; 2000US-192934P.  
18-MAY-2000; 2000US-205020P.  
18-MAY-2000; 2000US-205026P.  
19-MAY-2000; 2000US-205767P.  
19-MAY-2000; 2000US-205769P.  
(NOV5 ) NOVARTIS AG.  
(NOV5 ) NOVARTIS-ERFINDUNGEN VERW GES MBH.  
Werner G, Phares W, Jaritz M, Lapp H, Kalthoff FS;  
WPI: 2001-616466/71.  
P-PSDB; AAEI2080.  
New polypeptides for screening therapeutic agonists and antagonists comprise dendritic cell polypeptides -  
Claim 1; Page 51; 52pp; English.  
The invention relates to dendritic cell (DC) proteins and their corresponding DNA molecules. A pharmaceutical composition comprising agonist and antagonist of DC proteins are useful for treating abnormal conditions related to both an excess of and insufficient level of expression of DC gene, or related to both an excess of and insufficient activity of DC protein. Soluble form of DC proteins are used as an active ingredient in combination with pharmaceutical acceptable carriers. DC genes and proteins are useful for treating chronic inflammatory diseases, autoimmune diseases, transplant rejection crisis, including inflammatory skin diseases such as contact hypersensitivity, atopic dermatitis or virally-induced immune suppression such as AIDS and cancer. DC protein is useful for inducing immunological response in a mammal, and as immunogen to produce antibodies immunospecific for the polypeptide. DC gene is useful in gene therapy. DC gene is also useful as a diagnostic reagent, and for chromosomal identification. The present sequence is dendritic cell (DC) DCLEC/SPLICE 1 variant gene which is found to belong to the family of C-type lectins with one single carbohydrate recognition domain at the C-terminal end.  
Sequence 800 BP; 231 A; 160 C; 183 G; 226 T; 0 other;

Alignment Scores:  
Pred. No.: 2,05e-57 Length: 800  
Score: 617.00 Matches: 111  
Percent Similarity: 67.14% Conservative: 30

DT	XX	Best Local Similarity: 52.86%	Mismatches: 63
DE	XX	Query Match: 52.96%	Indels: 6
XX	XX	DB: 22	Gaps: 2
US-09-766-511b-53 (1-209) x AAD19730 (1-800)			
QY	6	GlnProGlnSerThrGluLysArg-----GlyTrpLeuSerLeuArgLeuTrpSer	22
DB	2	CGCCCGCGTCGCGAGACCGAGAGAAAGGACTCTGGTGGTTCAGTTGAAGTCTGGTCC	61
QY	23	ValAlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysValValThr	42
DB	62	ATGGCAGTCGTATCCACTTGTCTCAGTCTGTCTTTCACGTGTAGTCTCTGGTGGCT	121
QY	43	TyrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeu-----HisSer	59
DB	122	CACAATTTATGTATAGCAAACTGTCAAGAGCGCTGTCCAAGTTACGAGAGTATCAACAG	181
QY	60	TyrHisSerSerLeuThrCysPheSerGluGlyThrLysValProAlaTrpGlyCysCys	79
DB	182	TATCATCCAAAGCTGACCTCGCTCATGGAAGCAAGACATAGACATTGGAGCTGCTCC	241
QY	80	ProAlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGluGluLysVal	99
DB	242	CCAAACCCCTTGGACTTCATTTCAGTCTAGTTCTACTTATTTCTACTGGGATGCAATCT	301
QY	100	TrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsnThr	119
DB	302	TGGACTAAGAGTCAAAAGAACTGTCTGTGTGGGGGTGATCTGGTGGTGTATCAACACC	361
QY	120	GluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerPheLeuGly	139
DB	362	AGGGAAGAACAGGATTCATCATTCAGATCTGAAAAGAAATCTCTTATTTCTGGGG	421
QY	140	LeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrProTrpGluLys	159
DB	422	CTGTCAAGTCCAGGGGGTCGGCGACATTGGCAATGGTGTGACACACACATCAATGAA	481
QY	160	AsnValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAlaSerIle	179
DB	482	AATGTCACATTCCTGGCAGCTCAGGTGAACCAATAACCTTGATGCGGTGCGGATAATA	541
QY	180	ValPheTrpLysProThrGlyTrpAsnAspValIleCysGluThrArgArgAsn	199
DB	542	AATTTCCGTTCTTCAGAACAAATGGGGCTGGAATGATGATGATGATGATGATGATGAT	601
QY	200	SerIleCysGluMetAsnLysIleTyrLeu	209
DB	602	TCAATTTGCAAGATGAAGAAGATCTACATA	631
RESULT 11			
ABK52914	ID	ABK52914 standard; cDNA; 444 BP.	
XX	AC	ABK52914;	
XX	DT	27-AUG-2002 (first entry)	
XX	DE	Human cDNA encoding a partial dendritic cell immunoreceptor #2.	
XX	KW	Human; ss; gene; dendritic cell immunoreceptor; cytostatic;	
XX	KW	antiallergic; anorectic; antidiabetic; cancer; allergy; anaphylaxis;	
XX	KW	asthma; inflammation; obesity; diabetes; central nervous system disorder;	
XX	KW	Alzheimer's disease; Parkinson's disease; dementia; osteoarthritis;	
XX	KW	cardiovascular disorder; myocardial infarction; ischaemic heart disease;	
XX	KW	congestive heart failure; chronic obstructive pulmonary disease; COPD.	
XX	OS	Homo sapiens.	
XX	XX	Key Location/Qualifiers	
FT	CDS	1..444	
FT	FT	/*tag= a	
FT	FT	/product= "Dendritic cell immunoreceptor"	

```

FT /partial
XX /note= "No start or stop codon shown"
PN WO200232958-A2.
XX
XX
XX
PD 25-APR-2002.
XX
XX PF 12-OCT-2001; 2001WO-EP11812.
XX PR 16-OCT-2000; 2000US-240096P.
XX PR 27-AUG-2001; 2001US-314661P.
XX
XX
PA (FARB ) BAYER AG.
XX
XX Smolyar A;
XX WPI; 2002-463308/49.
XX P-PSDB; AAU98015.
XX
XX Novel human dendritic cell immunoreceptor polypeptide, useful for
XX treating cancer, asthma, obesity, diabetes, central nervous system
XX disorder or cardiovascular disorder
XX
XX Claim 1; Fig 15; 114pp; English.
XX
XX The invention relates to a purified human dendritic cell immunoreceptor
XX polypeptide and its encoding nucleic acid. Also included are an
XX expression vector comprising the nucleic acid, a host cell containing
XX the vector, reducing the activity of human dendritic cell immunoreceptor
XX involves contacting a cell with a reagent which specifically binds to
XX the polypeptide or nucleic acid and a reagent that modulates the activity
XX of the polypeptide or nucleic acid. The reagent is useful for the
XX preparation of a medicament for modulating the activity of human
XX dendritic cell immunoreceptor in a disease such as cancer, allergy,
XX anaphylaxis, asthma, inflammation, obesity, diabetes, a central nervous
XX system (CNS) disorder (e.g. Alzheimer's disease, Parkinson's disease or
XX dementia), or a cardiovascular disorder (e.g. myocardial infarction,
XX ischaemic heart disease, congestive heart failure), chronic obstructive
XX pulmonary disease (COPD) and osteoarthritis (many other diseases and
XX disorders are listed in the specification). The polypeptide and nucleic
XX acid are useful for identifying test compounds which act as agonists or
XX antagonists, for raising specific antibodies, and as a bait protein in a
XX two-hybrid or three-hybrid assay. The nucleic acid is useful in
XX diagnostic assays for detecting diseases and abnormalities or
XX susceptibility to disease and abnormalities related to the presence of
XX mutations. The present sequence is a cDNA encoding a partial
XX dendritic cell immunoreceptor.
XX
XX SQ Sequence 444 BP; 123 A; 86 C; 101 G; 134 T; 0 other;

Alignment Scores:
Pred. No.: 1.15e-40 Length: 444
Score: 459.50 Matches: 80
Percent Similarity: 67.57% Conservative: 20
Best Local Similarity: 54.05% Mismatches: 47
Query Match: 39.44% Indels: 1
DB: 24 Gaps: 1

US-09-766-511B-53 (1-209) x ABK52914 (1-444)
QY 63 SerLeuThrCysPheSerGluGlyThrLysVal---ProAlaTrpGlyCysCysProAla 81
Db 1 TCTGTAAATATTTCCTCTCTATACTCAATTTGGGAAGATTGGAGCTGCGCCACACC 60
QY 82 SerTrpLysSerPheGlySerCysTyrPheIleSerSerGluGluLysValTrpSer 101
Db 61 CCTTGGACTCAATTCAGTCTAGTTCCTCTCTATACTCAATTTGGGAAGATTGGAGCTGCGCCACACC 120
QY 102 LysSerGluGlnAsnGlyValGluMetGlyAlaHisLeuValPheAsnThrGluAla 121
Db 121 AGAGTCAAAAGAACTGTTCTGTGATGGGCTGATCTGGTGGTGCATCAACACACGGGAA 180
QY 122 GluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeuSer 141

```

```

Db 181 GAACAGGATTTCATCAATTCAGAAATCGAAAGAAATCTCTTATTTCTGGGGCTGCA 240
QY 142 AspProGlnGlnAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsnVal 161
Db 241 GATCCAGGGGTCGGCGACATTGGCAATGGTTGACCAGACCATCAATCAATAATGTC 300
QY 162 ArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAlaSerIleValPhe 181
Db 301 ACATTCTGGCACTCAGTGAACCAATAACCTTGATGAGCGTTGTGCGATAATAAATTC 360
QY 182 TrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArgAsnSerIle 201
Db 361 CGTTCTTCAGAGAATGGGCTGGAATGACATTCATCTGTCATGCTCAGAGTCAATT 420
QY 202 CysGluMetAsnLysIleTyrLeu 209
Db 421 TGCAGATGAAGAAGATCTACATA 444

RESULT 12
ABK52903
ID ABK52903 standard; cDNA; 402 BP.
XX
XX AC ABK52903;
XX
XX DT 27-AUG-2002 (first entry)
XX DE Human cDNA encoding a partial dendritic cell immunoreceptor #1.
XX KW Human; ss; gene; dendritic cell immunoreceptor; cytostatic;
XX KW antasthmatic; anorectic; antidiabetic; cancer; allergy; anaphylaxis;
XX KW asthma; inflammation; obesity; diabetes; central nervous system disorder;
XX KW Alzheimer's disease; Parkinson's disease; dementia; osteoarthritis;
XX KW cardiovascular disorder; myocardial infarction; ischaemic heart disease;
XX KW congestive heart failure; chronic obstructive pulmonary disease; COPD.
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX CDS 1..402
XX FT /*tag= a
XX FT /product= "Dendritic cell immunoreceptor"
XX FT /partial
XX FT /note= "No start or stop codon shown"
XX
XX PN WO200232958-A2.
XX
XX PD 25-APR-2002.
XX
XX PF 12-OCT-2001; 2001WO-EP11812.
XX
XX PR 16-OCT-2000; 2000US-240096P.
XX PR 27-AUG-2001; 2001US-314661P.
XX
XX PA (FARB ) BAYER AG.
XX
XX PI Smolyar A;
XX
XX DR WPI; 2002-463308/49.
XX DR P-PSDB; AAU98013.
XX
XX PT Novel human dendritic cell immunoreceptor polypeptide, useful for
XX PT treating cancer, asthma, obesity, diabetes, central nervous system
XX PT disorder or cardiovascular disorder
XX
XX PS Claim 1; Fig 1; 114pp; English.
XX
XX The invention relates to a purified human dendritic cell immunoreceptor
XX polypeptide and its encoding nucleic acid. Also included are an
XX expression vector comprising the nucleic acid, a host cell containing
XX the vector, reducing the activity of human dendritic cell immunoreceptor
XX involves contacting a cell with a reagent which specifically binds to
XX the polypeptide or nucleic acid and a reagent that modulates the activity
XX of the polypeptide or nucleic acid. The reagent is useful for the
XX preparation of a medicament for modulating the activity of human
XX dendritic cell immunoreceptor in a disease such as cancer, allergy,
XX anaphylaxis, asthma, inflammation, obesity, diabetes, a central nervous
XX system (CNS) disorder (e.g. Alzheimer's disease, Parkinson's disease or
XX dementia), or a cardiovascular disorder (e.g. myocardial infarction,
XX ischaemic heart disease, congestive heart failure), chronic obstructive
XX pulmonary disease (COPD) and osteoarthritis (many other diseases and
XX disorders are listed in the specification). The polypeptide and nucleic
XX acid are useful for identifying test compounds which act as agonists or
XX antagonists, for raising specific antibodies, and as a bait protein in a
XX two-hybrid or three-hybrid assay. The nucleic acid is useful in
XX diagnostic assays for detecting diseases and abnormalities or
XX susceptibility to disease and abnormalities related to the presence of
XX mutations. The present sequence is a cDNA encoding a partial
XX dendritic cell immunoreceptor.
XX
XX SQ Sequence 444 BP; 123 A; 86 C; 101 G; 134 T; 0 other;

```

CC of the polypeptide or nucleic acid. The reagent is useful for the  
 CC preparation of a medicament for modulating the activity of human  
 CC dendritic cell immunoreceptor in a disease such as cancer, allergy,  
 CC anaphylaxis, asthma, inflammation, obesity, diabetes, a central nervous  
 CC system (CNS) disorder (e.g. Alzheimer's disease, Parkinson's disease or  
 CC dementia), or a cardiovascular disorder (e.g. myocardial infarction,  
 CC ischaemic heart disease, congestive heart failure), chronic obstructive  
 CC pulmonary disease (COPD) and osteoarthritis (many other diseases and  
 CC disorders are listed in the specification). The polypeptide and nucleic  
 CC acid are useful for identifying test compounds which act as agonists or  
 CC antagonists, for raising specific antibodies, and as a bait protein in a  
 CC two-hybrid or three-hybrid assay. The nucleic acid is useful in  
 CC diagnostic assays for detecting diseases and abnormalities or  
 CC susceptibility to disease and abnormalities related to the presence of  
 CC mutations. The present sequence is a cDNA encoding a partial  
 CC dendritic cell immunoreceptor.

XX Sequence 402 BP; 114 A; 80 C; 96 G; 112 T; 0 other;

Alignment Scores:  
 Pred. No.: 1.13e-40 Length: 402  
 Score: 459.00 Matches: 77  
 Percent Similarity: 70.90% Conservative: 18  
 Best Local Similarity: 57.46% Mismatches: 39  
 Query Match: 39.40% Indels: 0  
 DB: 24 Gaps: 0

US-09-766-511B-53 (1-209) x ABK52903 (1-402)

Qy 76 TrpGlyCysCysProAlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSer 95  
 Db 1 TGGAGCTGCTGCCAACCCCTTGGACTTCATTTTCAGTCTAGTTCCTACTTTTCTACT 60  
 Qy 96 GluGluLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuVal 115  
 Db 61 GGGATGCAATCTTGGACTAAGAGTCAAAGAACTGTTCTGTGATGGGGCTGATCTGGTG 120  
 Qy 116 ValPheAsnThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSer 135  
 Db 121 GTGATCAACACCGGGAGAGACAGATTTTCATCTCAGATTCGAAAGAAATCTTCT 180  
 Qy 136 TyrPheLeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThr 155  
 Db 181 TATTTCTGGGCTGTCCAGATCCAGGGGTCGGCGACATTTGGCAATGGTTGACACACA 240  
 Qy 156 ProTyrGluLysAsnValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGln 175  
 Db 241 CCATACATGAATAATGTACATTTCTGGCACTTCAGGTGAACCCAAATAACCTTGATGACGT 300  
 Qy 176 CysAlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGlu 195  
 Db 301 TGTGGATAATAAATTTCCGTTCTTCAGAGAATGGGGCTGGGAATGACATTCATGTCAT 360  
 Qy 196 ThrArgArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209  
 Db 361 GTACTCAGAAGTCAATTTGCAAGATGAAGAAGATCTACATA 402

RESULT 13  
 AAS31385  
 ID AAS31385 standard; cDNA; 1091 BP.  
 XX  
 AC AAS31385;  
 XX  
 DT 04-DEC-2001 (first entry)  
 DE Human cDNA encoding a novel extracellular matrix protein, Seq ID No 199.  
 XX  
 KW Human: secreted extracellular matrix protein; ss: immunomodulatory;  
 KW Anti-HIV; antianemic; antirheumatic; antisclerotic; cardiac; vascular;  
 KW cerebroprotective; thrombolytic; antimicrobial; ophthalmic; cytostatic;  
 KW antiaizheimers; immune/autoimmune disease; HIV infection; anaemia;  
 KW human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;  
 KW cancers; hyperproliferative disorder; breast neoplasm; melanoma;

KW Sezary syndrome; Gaucher's disease; neurological diseases;  
 KW Alzheimer's disease; Parkinson's disease; cardiovascular disorder;  
 KW cardiac arrest; tachycardia; angina; infection; corneal infections;  
 KW wound healing; immunogen; gene therapy; antisense; food additive.  
 XX Homo sapiens.  
 XX WO200155368-A1.  
 XX 02-AUG-2001.  
 XX 17-JAN-2001; 2001WO-US01348.  
 PF 31-JAN-2000; 2000US-0179065.  
 PR 04-FEB-2000; 2000US-0180628.  
 PR 24-FEB-2000; 2000US-0184664.  
 PR 02-MAR-2000; 2000US-0186350.  
 PR 16-MAR-2000; 2000US-0189874.  
 PR 17-MAR-2000; 2000US-0190076.  
 PR 18-APR-2000; 2000US-0198123.  
 PR 19-MAY-2000; 2000US-0205515.  
 PR 07-JUN-2000; 2000US-0209467.  
 PR 28-JUN-2000; 2000US-0214886.  
 PR 30-JUN-2000; 2000US-0215135.  
 PR 07-JUL-2000; 2000US-0216647.  
 PR 07-JUL-2000; 2000US-0216880.  
 PR 11-JUL-2000; 2000US-0217487.  
 PR 11-JUL-2000; 2000US-0217496.  
 PR 16-JUL-2000; 2000US-0218290.  
 PR 26-JUL-2000; 2000US-0220963.  
 PR 26-JUL-2000; 2000US-0220964.  
 PR 14-AUG-2000; 2000US-0224518.  
 PR 14-AUG-2000; 2000US-0224519.  
 PR 14-AUG-2000; 2000US-0225213.  
 PR 14-AUG-2000; 2000US-0225214.  
 PR 14-AUG-2000; 2000US-0225266.  
 PR 14-AUG-2000; 2000US-0225267.  
 PR 14-AUG-2000; 2000US-0225268.  
 PR 14-AUG-2000; 2000US-0225270.  
 PR 14-AUG-2000; 2000US-0225447.  
 PR 14-AUG-2000; 2000US-0225757.  
 PR 14-AUG-2000; 2000US-0225758.  
 PR 14-AUG-2000; 2000US-0225759.  
 PR 18-AUG-2000; 2000US-0226279.  
 PR 22-AUG-2000; 2000US-0226681.  
 PR 22-AUG-2000; 2000US-0226681.  
 PR 22-AUG-2000; 2000US-0226868.  
 PR 22-AUG-2000; 2000US-0227182.  
 PR 23-AUG-2000; 2000US-0227009.  
 PR 30-AUG-2000; 2000US-0228924.  
 PR 01-SEP-2000; 2000US-0229287.  
 PR 01-SEP-2000; 2000US-0229343.  
 PR 01-SEP-2000; 2000US-0229344.  
 PR 01-SEP-2000; 2000US-0229345.  
 PR 05-SEP-2000; 2000US-0229509.  
 PR 05-SEP-2000; 2000US-0229513.  
 PR 06-SEP-2000; 2000US-0230437.  
 PR 06-SEP-2000; 2000US-0230438.  
 PR 08-SEP-2000; 2000US-0231242.  
 PR 08-SEP-2000; 2000US-0231243.  
 PR 08-SEP-2000; 2000US-0231244.  
 PR 08-SEP-2000; 2000US-0231413.  
 PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232080.  
 PR 08-SEP-2000; 2000US-0232081.  
 PR 12-SEP-2000; 2000US-0231968.  
 PR 14-SEP-2000; 2000US-0232397.  
 PR 14-SEP-2000; 2000US-0232398.  
 PR 14-SEP-2000; 2000US-0232399.  
 PR 14-SEP-2000; 2000US-0232400.  
 PR 14-SEP-2000; 2000US-0232401.  
 PR 14-SEP-2000; 2000US-0233063.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 14-SEP-2000; 2000US-0233065.

PR 21-SEP-2000; 2000US-0234223.  
 PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236602.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 17-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 08-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.

PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX Rosen CA, Barash SC, Ruben SM;  
 PI P-PSDB; RAU19814.  
 XX WPI; 2001-465572/50.  
 DR Nucleic acid molecules encoding human secreted extracellular matrix  
 XX proteins, used in preventing, treating or ameliorating a disorder, e.g.  
 PT Alzheimer's and Parkinson's diseases and cancers -  
 PT  
 XX Claim 1; SEQ ID No 199; 577pp; English.  
 PS  
 XX The invention relates to isolated nucleic acid molecules encoding  
 CC novel human secreted extracellular matrix proteins (SPs). The  
 CC polynucleotides and proteins are used to prevent, treat a medical  
 CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,  
 CC chickens or sheep. For example, disorders associated with decreased  
 CC expression of SPs. The SP polynucleotide or a vector expressing them may  
 CC be administered to treat diseases by gene therapy. Antisense molecules  
 CC may be administered to down regulate expression of SPs by binding with  
 CC the cells own genes and preventing their expression. The polynucleotides  
 CC may also be used as DNA probes in diagnostic assays. The SPs may also be  
 CC used as antigens to produce antibodies and to identify modulators  
 CC (agonists and antagonists) of the SPs. The anti-SP antibodies and  
 CC antagonists may also be used to down regulate expression and activity of  
 CC SP and as diagnostic agents for detecting the presence of SPs in samples.  
 CC The disorders include for example: immune/autoimmune diseases (e.g. HIV  
 CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis  
 CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.  
 CC melanomas, neoplasms of the breast or liver, Sezary syndrome and  
 CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,  
 CC Parkinson's disease) cardio-/cerebrovascular disorders (e.g. cardiac  
 CC arrest, tachycardia and angina), infections caused by bacteria, viruses  
 CC and fungi and ocular disorders (e.g. corneal infections). Other uses  
 CC include wound healing, maintenance of organs before transplantation,  
 Alignment Scores:  
 Pred. No.: 6.16e-40 Length: 1091  
 Score: 458.00 Matches: 83  
 Percent Similarity: 62.57% Conservative: 34  
 Best Local Similarity: 44.39% Mismatches: 66  
 Query Match: 39.31% Indels: 4  
 DB: 22 Gaps: 3  
 US-09-766-511b-53 (1-209) x AAS31385 (1-1091)  
 QY 26 IleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysValValThrTyrHisPhe 45  
 Db 172 ATATTTTCTGCTATTTGGCAATCTCATTTCTTTTGTTCATTTCCTTCAA--- 228  
 QY 46 ThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHisSerTyrHisSerSerLeuThr 65  
 Db 229 AAATATTTCTCAGCTTCTTGAAAAAAGACTACAAAGAGCTGGTTCATACAAATTGGAG 288  
 QY 66 CysPheSerGluGlyThrLysVal-----ProAlaTrpGlyCysCysProAlaSerTrp 83  
 Db 289 TGTGTGAAAAAATAATGCCCGTGGAGAGACAGCCCTGGAGCTGTGCCCAAGAAATTGG 348  
 QY 84 LysSerPheGlySerSerCysTyrPheIleSerSerGluGluLysValTrpSerLysSer 103  
 Db 349 AAGTCATTTCCTCCACTGCTACTTTATTTCTACAGTACAGCATCTTGGCAAGACAGT 408  
 QY 104 GluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsnThrGluAlaGluIn 123  
 Db 409 GAGAAAGGACTGTGCTAGATGGAGGCTCACCTGCTGTGATATAAACAACACTCAAGAAGAGCAG 468  
 QY 124 AsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPheGluGlyLeuSerAspPro 143



Db 469 GATTTCATCTTCAGAAATCGCAAGAAGAAATCTGCTATTATTTGTGGGCTCTCATGATCCA 528  
QY 144 GlnGlyAsnAsnAsnTrrpGlnTrrpIleAspLysThrProTyrGluLysAsnValArgPhe 163  
Db 529 GAAGGTTCAGGACATTGGCAATGGTGTGATCAGACACCATACAATGAAGATTCACATTC 588  
QY 164 TrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAlaSerIleValPheTrpLys 183  
Db 589 TGGCATCCACGTGAGCCAGTGCATCCCAATGAGCGTGGCTGTGCTAAATTTTCGTAAA 648  
QY 184 ---ProThrClyTrpGlyTrpAsnAspValIleCysGluThrArgArgAsnSerIleCys 202  
Db 649 TCACCAAAAGATGGGCGTGAATGATTAATGTCTGGTCCCTCAAGGTCAGTTGT 708  
QY 203 GluMetAsnLysIleTyrLeu 209  
Db 709 GAGATGATGAAGATCCACTTA 729  
RESULT 14  
ABQ66709  
ID ABQ66709 standard; cdNA; 1091 BP.  
XX AC ABQ66709;  
XX XX  
DT 23-AUG-2002 (first entry)  
XX XX  
DE Human polynucleotide SEQ ID NO 199.  
XX XX  
KW Human; nootropic; neuroprotective; cytostatic; dermatologic; virocidic;  
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;  
KW antiparkinsonian; antispasmodic; antianemic; antiarthritic; cancer;  
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;  
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;  
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; nephrotropic; gene therapy; vaccine;  
KW gene; ss.  
XX XX  
OS Homo sapiens.  
XX XX  
PN US2002042386-A1.  
XX XX  
PD 11-APR-2002.  
XX XX  
PF 17-JAN-2001; 2001US-0764870.  
XX XX  
PR 31-JAN-2000; 2000US-179065P.  
PR 04-FEB-2000; 2000US-180628P.  
PR 28-JUN-2000; 2000US-214886P.  
PR 07-JUL-2000; 2000US-216647P.  
PR 07-JUL-2000; 2000US-216880P.  
PR 11-JUL-2000; 2000US-217487P.  
PR 11-JUL-2000; 2000US-217496P.  
PR 14-JUL-2000; 2000US-218290P.  
PR 26-JUL-2000; 2000US-220963P.  
PR 26-JUL-2000; 2000US-220964P.  
PR 14-AUG-2000; 2000US-224518P.  
PR 14-AUG-2000; 2000US-224519P.  
PR 14-AUG-2000; 2000US-225267P.  
PR 14-AUG-2000; 2000US-225268P.  
PR 14-AUG-2000; 2000US-225270P.  
PR 14-AUG-2000; 2000US-225447P.  
PR 14-AUG-2000; 2000US-225757P.  
PR 14-AUG-2000; 2000US-225758P.  
PR 22-AUG-2000; 2000US-226868P.  
PR 30-AUG-2000; 2000US-228924P.  
PR 01-SEP-2000; 2000US-228928P.  
PR 01-SEP-2000; 2000US-229343P.  
PR 01-SEP-2000; 2000US-229344P.  
PR 01-SEP-2000; 2000US-229345P.  
PR 05-SEP-2000; 2000US-229509P.  
PR 05-SEP-2000; 2000US-229513P.  
PR 08-SEP-2000; 2000US-231413P.  
PR 21-SEP-2000; 2000US-234223P.

PR 21-SEP-2000; 2000US-234274P.  
PR 25-SEP-2000; 2000US-234997P.  
PR 27-SEP-2000; 2000US-235834P.  
PR 29-SEP-2000; 2000US-236327P.  
PR 29-SEP-2000; 2000US-236367P.  
PR 29-SEP-2000; 2000US-236368P.  
PR 29-SEP-2000; 2000US-236369P.  
PR 29-SEP-2000; 2000US-236370P.  
PR 02-OCT-2000; 2000US-236802P.  
PR 02-OCT-2000; 2000US-237037P.  
PR 02-OCT-2000; 2000US-237038P.  
PR 02-OCT-2000; 2000US-237039P.  
PR 02-OCT-2000; 2000US-237040P.  
PR 13-OCT-2000; 2000US-239935P.  
PR 20-OCT-2000; 2000US-240960P.  
PR 20-OCT-2000; 2000US-241785P.  
PR 20-OCT-2000; 2000US-241809P.  
PR 01-NOV-2000; 2000US-244617P.  
PR 17-NOV-2000; 2000US-249299P.  
PR 08-DEC-2000; 2000US-251856P.  
PR 08-DEC-2000; 2000US-251868P.  
PR 08-DEC-2000; 2000US-251869P.  
XX XX  
PA (ROSE/) ROSEN C A.  
PA (RUBE/) RUBEN S M.  
PA (BARA/) BARASH S C.  
XX XX  
PI Rosen CA, Ruben SM, Barash SC;  
XX XX  
DR WPI; 2002-470713/50.  
DR P-PSDB; ABP48034.  
XX XX  
PT New nucleic acid encoding human proteins, useful for diagnosis,  
PT treatment and prevention of e.g. osteoporosis, also related  
PT polypeptides and antibodies  
XX XX  
PS Claim 1; SEQ ID NO 199; 235pp + Sequence Listing; English.  
XX XX  
CC The invention relates to novel genes (ABQ66521-ABQ66785) and proteins  
CC (ABP47846-ABP48110) useful for preventing, treating or ameliorating  
CC medical conditions e.g. by protein or gene therapy. The genes are  
CC isolated from a range of human tissues disclosed in the specification.  
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful  
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone  
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;  
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
CC and parasitic infections.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from USPTO at seqdata.uspto.gov/sequence.html?docid=999909764870.  
XX XX  
SQ Sequence 1091 BP; 308 A; 213 C; 250 G; 320 T; 0 other;

Alignment Scores:  
Pred. No.: 6.16e-40 Length: 1091  
Score: 458.00 Matches: 83  
Percent Similarity: 62.57% Conservative: 34  
Best Local Similarity: 44.39% Mismatches: 66  
Query Match: 39.31% Indels: 4  
DB: 24 Gaps: 3

US-09-766-511B-53 (1-209) x ABQ66709 (1-1091)

QY 26 IleserIleAlaLeuLeuSerAlaCysPheIleValserCysValValThrTyrHisPhe 45

Db 172 ATATTTTCTCTGATATGGCAATCTCATCTTTATTCCTTTCTTTCTTTCTTTCAA--- 228

Qy 46 ThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHisSerTyrHisSerSerLeuThr 65  
Db 229 AAATATTCTCAGCTTCTTGAATAAAGACTACAAAGAGCTGGTTTCATACAACTGGAG 288  
Qy 66 CysPheSerGluGlyThrLysVal-----ProAlaTrpGlyCysCysProAlaSerTrp 83  
Db 289 TGTGTGAAAAAATATGCGCGTGAAGAGACACCGCTGGAGCTGTTCGCCAAAGAATGG 348  
Qy 84 LysSerPheGlySerSerCysTyrPheIleSerSerGluGluLysValTrpSerLysSer 103  
Db 349 AAGTCATTAGTTCCTCACTGCTACTTTATTCTACTCAATCAGCATCTTGGCAAGACAGT 408  
Qy 104 GluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsnThrGluAlaGluGln 123  
Db 409 GAGAAGGACTGTGTAGAAATGGAGGCTCAGCTGTGTGTGATAAACACACTCAAGAAGACAG 468  
Qy 124 AsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeuSerAspPro 143  
Db 469 GATTTCATCTCCAGATCTGCAGAAGATCTGCTTATTATTGTGGGCTCTCAGATCCA 528  
Qy 144 GlnGlyAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsnValArgPhe 163  
Db 529 GAAGGTCAGGCACATTGGCAATGGTTGATCAGACACCATACATGAAGTTCACCATTC 588  
Qy 164 TrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAlaSerIleValPheTrpLys 183  
Db 589 TGGCATCCAGTGAGCCCATGATCCCAATGAGCGTGGCTGTGTGATAATTTCGTAAA 648  
Qy 184 ---ProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgAsnSerIleCys 202  
Db 649 TCACCCAAAGATGGGCTGGAATGATGTTAATTGCTTGGTCTCTCAAGGTCAGTTGT 708  
Qy 203 GluMetAsnLysIleTyrLeu 209  
Db 709 GAGATGATGAAGATCCACTTA 729  
RESULT 15  
AAS31224  
ID AAS31224 standard; cDNA; 1096 bp.  
XX AAS31224;  
AC AAS31224;  
DT 04-DEC-2001 (first entry)  
XX Human cDNA encoding a novel extracellular matrix protein, Seq ID No 38.  
XX Human; secreted extracellular matrix protein; ss; immunomodulatory;  
KW Anti-HIV; antianemic; antirheumatic; antisclerotic; cardiac; vascular;  
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmic; cytostatic;  
KW antialzheimers; immune/autoimmune disease; HIV infection; anaemia;  
KW human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;  
KW cancers; hyperproliferative disorder; breast neoplasm; melanoma;  
KW Sezary syndrome; Gaucher's disease; neurological diseases;  
KW Alzheimer's disease; Parkinson's disease; cardiovascular disorder;  
KW cardiac arrest; tachycardia; angina; infection; corneal infections;  
KW wound healing; immunogen; gene therapy; antisense; food additive.  
XX Homo sapiens.  
OS  
XX  
PN WO20015368-A1.  
XX  
PD 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01348.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226688.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 05-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 21-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 29-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.

PR 20-OCT-2000; 2000US-0241785;  
 PR 20-OCT-2000; 2000US-0241786;  
 PR 20-OCT-2000; 2000US-0241787;  
 PR 20-OCT-2000; 2000US-0241808;  
 PR 20-OCT-2000; 2000US-0241809;  
 PR 01-NOV-2000; 2000US-0241826;  
 PR 2000US-0244617;  
 PR 2000US-0244674;  
 PR 2000US-0246475;  
 PR 08-NOV-2000; 2000US-0246476;  
 PR 08-NOV-2000; 2000US-0246477;  
 PR 08-NOV-2000; 2000US-0246478;  
 PR 08-NOV-2000; 2000US-0246523;  
 PR 08-NOV-2000; 2000US-0246524;  
 PR 08-NOV-2000; 2000US-0246525;  
 PR 08-NOV-2000; 2000US-0246526;  
 PR 08-NOV-2000; 2000US-0246527;  
 PR 08-NOV-2000; 2000US-0246528;  
 PR 08-NOV-2000; 2000US-0246532;  
 PR 08-NOV-2000; 2000US-0246609;  
 PR 08-NOV-2000; 2000US-0246610;  
 PR 08-NOV-2000; 2000US-0246611;  
 PR 08-NOV-2000; 2000US-0246613;  
 PR 17-NOV-2000; 2000US-0249207;  
 PR 17-NOV-2000; 2000US-0249208;  
 PR 17-NOV-2000; 2000US-0249209;  
 PR 17-NOV-2000; 2000US-0249210;  
 PR 17-NOV-2000; 2000US-0249211;  
 PR 17-NOV-2000; 2000US-0249212;  
 PR 17-NOV-2000; 2000US-0249213;  
 PR 17-NOV-2000; 2000US-0249214;  
 PR 17-NOV-2000; 2000US-0249215;  
 PR 17-NOV-2000; 2000US-0249216;  
 PR 17-NOV-2000; 2000US-0249217;  
 PR 17-NOV-2000; 2000US-0249218;  
 PR 17-NOV-2000; 2000US-0249244;  
 PR 17-NOV-2000; 2000US-0249245;  
 PR 17-NOV-2000; 2000US-0249264;  
 PR 17-NOV-2000; 2000US-0249265;  
 PR 17-NOV-2000; 2000US-0249266;  
 PR 17-NOV-2000; 2000US-0249297;  
 PR 17-NOV-2000; 2000US-0249299;  
 PR 17-NOV-2000; 2000US-0249300;  
 PR 01-DEC-2000; 2000US-0250160;  
 PR 01-DEC-2000; 2000US-0250391;  
 PR 05-DEC-2000; 2000US-0251030;  
 PR 05-DEC-2000; 2000US-0251988;  
 PR 05-DEC-2000; 2000US-0256719;  
 PR 06-DEC-2000; 2000US-0251479;  
 PR 08-DEC-2000; 2000US-0251856;  
 PR 08-DEC-2000; 2000US-0251868;  
 PR 08-DEC-2000; 2000US-0251869;  
 PR 08-DEC-2000; 2000US-0251989;  
 PR 08-DEC-2000; 2000US-0251990;  
 PR 11-DEC-2000; 2000US-0254097;  
 PR 05-JAN-2001; 2001US-0259678;  
 (HUMA-) HUMAN GENOME SCI INC.  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX WPI; 2001-465572/50.  
 DR P-PSDB; AAU19653.  
 XX  
 PT Nucleic acid molecules encoding human secreted extracellular matrix  
 PT proteins, used in preventing, treating or ameliorating a disorder, e.g.  
 PT Alzheimer's and Parkinson's diseases and cancers -  
 XX  
 PS Claim 1; SEQ ID No 38; 577pp; English.  
 XX  
 CC The invention relates to isolated nucleic acid molecules encoding  
 CC novel human secreted extracellular matrix proteins (SPs). The  
 CC polynucleotides and proteins are used to prevent, treat a medical  
 CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,

CC chickens or sheep. For example, disorders associated with decreased  
 CC expression of SPs. The SP polynucleotide or a vector expressing them may  
 CC be administered to treat diseases by gene therapy. Antisense molecules  
 CC may be administered to down regulate expression of SPs by binding with  
 CC the cells own genes and preventing their expression. The polynucleotides  
 CC may also be used as DNA probes in diagnostic assays. The SPs may also be  
 CC used as antigens to produce antibodies and to identify modulators  
 CC (agonists and antagonists) of the SPs. The anti-(SP) antibodies and  
 CC antagonists may also be used to down regulate the presence and activity of  
 CC SP and as diagnostic agents for detecting the presence of SPs in samples.  
 CC The disorders include for example: immune/autoimmune diseases (e.g. HIV  
 CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis  
 CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.  
 CC melanomas, neoplasms of the breast or liver, Sezary syndrome and  
 CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,  
 CC Parkinson's disease) cardio-/cerebrovascular disorders (e.g. cardiac  
 CC arrest, tachycardia and angina), infections caused by bacteria, viruses  
 CC and fungi and ocular disorders (e.g. corneal infections). Other uses  
 CC include wound healing, maintenance of organs before transplantation,

Alignment Scores:  
 Pred. No.: 6.2e-40 Length: 1096  
 Score: 458.00 Matches: 83  
 Percent Similarity: 62.57% Conservative: 34  
 Best Local Similarity: 44.39% Mismatches: 66  
 Query Match: 39.31% Indels: 4  
 DB: 22 Gaps: 3

US-09-766-511B-53 (1-209) x AAS31224 (1-1096)

Qy 26 IleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysValValThrTyrHisPhe 45  
 Db 182 ATATTTTCTCTGCTATTGGCAATCTCATTTCTTTATTTGTCATTTCTTTCAA--- 238  
 Qy 46 ThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHisSerTyrHisSerSerLeuThr 65  
 Db 239 AAATATTTCTAGCTTCTTGAAGAAAGACTACAAAGAGCTGGTTCATACAAATTGGAG 298  
 Qy 66 CysPheSerGluGlyThrLysVal-----ProAlaTrpGlyCysCysProAlaSerTrp 83  
 Db 299 TGTGTGAAATAAATATATCCCGTGGAGAGACAGCTGGAGCTGTGGCCCAAGAATTGG 358  
 Qy 84 LysSerPheGlySerCysTyrPheIleSerSerGluGluLysValTrpSerLysSer 103  
 Db 359 AAGTCATTTAGTTCACACTGCTACTTTATTTCTAGTAATCAGCATCTTGGCAAGACAGT 418  
 Qy 104 GluGlnAsnCysValGluMetGlyAlaHisLeuValPheAsnThrGluAlaGluGln 123  
 Db 419 GAGAAGGACTGTGTAGAAATGGAGGCTCACCTGCTGGTGAATAACACTCAAGAGAGCAG 478  
 Qy 124 AsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeuSerAspPro 143  
 Db 479 GATTTTCATCTCCAGAATCTGCAAGAAGAATCTGCTTATTTGTGGGCTCTCAGATCCA 538  
 Qy 144 GlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsnValArgPhe 163  
 Db 539 GAAGGTCAGCGACATTGGCAATGGTGTGATCAGACACCATACAAATGAAGTTCACATTC 598  
 Qy 164 TrpHisLeuGlyGluProAsnHisSerAlaGlnGlnCysAlaSerIleValPheTrpLys 183  
 Db 599 TGGCATCCAGCTGAGGCCAGTCATCCCAATGAGCCCTGCTGTGCTAAATTTTCGTAAA 658  
 Qy 184 ---ProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArgAsnSerIleCys 202  
 Db 659 TCACCCCAAGATGGGCTGGAATGATGCTTAATTGCTTGGTCCCTCAAAGGTCAGTTGT 718  
 Qy 203 GluMetAsnLysIleTyrLeu 209  
 Db 719 GAGATGATGAAGATCCACTTA 739

Search completed: February 20, 2003, 19:33:30  
 Job time : 224 secs



GenCore version 5.1.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: February 20, 2003, 18:37:42 : Search time 2230 Seconds  
(without alignments)  
2727.572 Million cell updates/sec

Title: US-09-766-511b-53  
Perfect score: 1165  
Sequence: 1 MMQEQPQSKRRKGLSURL.....NDVICETRRNSICENMKIYL 209

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09766511/runat\_19022003\_094338\_29286/app.query.fasta\_1.391  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USPR=US09766511.ecgn\_1.1.1616 @runat\_19022003\_094338\_29286 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:\*  
1: gb.ba:\*  
2: gb.htg:\*  
3: gb.in:\*  
4: gb.om:\*  
5: gb.ov:\*  
6: gb.pat:\*  
7: gb.ph:\*  
8: gb.pl:\*  
9: gb.pr:\*  
10: gb.ro:\*  
11: gb.sts:\*  
12: gb.sy:\*  
13: gb.un:\*  
14: gb.vi:\*  
15: em.ba:\*  
16: em.fun:\*  
17: em.hum:\*  
18: em.in:\*  
19: em.mu:\*  
20: em.om:\*  
21: em.or:\*  
22: em.ov:\*  
23: em.pat:\*  
24: em.ph:\*  
25: em.pl:\*  
26: em.ro:\*  
27: em.sts:\*  
28: em.un:\*

29: em.vi:\*  
30: em.htg\_hum:\*  
31: em.htg\_inv:\*  
32: em.htg\_other:\*  
33: em.htg\_mus:\*  
34: em.htg\_pln:\*  
35: em.htg\_rod:\*  
36: em.htg\_mam:\*  
37: em.htg\_vrt:\*  
38: em.sy:\*  
39: em.htgo\_hum:\*  
40: em.htgo\_mus:\*  
41: em.htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Query Length	DB ID	Description
1	793	68.1	1227	6	AX155225 Sequence
2	793	68.1	1227	10	AF240357 Mus muscu
3	793	68.1	1254	10	BC023008 Mus muscu
4	659	56.6	1125	10	AF240358 Mus muscu
5	637	54.7	851	9	AF325459 Homo sapi
6	637	54.7	1312	6	AX155223 Sequence
7	637	54.7	1313	9	AF293615 Homo sapi
8	630	54.1	827	6	AX357481 Sequence
9	617	53.0	800	6	AX357483 Sequence
10	592.5	50.9	1125	10	AF240359 Mus muscu
11	534.5	45.9	739	9	AF325460 Homo sapi
12	459.5	39.4	444	6	AX456975 Sequence
13	459	39.4	402	6	AX456962 Sequence
14	458	39.3	1133	9	AF109146 Homo sapi
15	458	39.3	1271	9	HSAL133532 Sequence
16	453	38.9	1305	9	AF067800 Homo sapi
17	445.5	38.2	1195	6	AX400064 Sequence
18	445	38.2	1308	9	AF328684 Homo sapi
19	445	38.2	1973	9	BC032313 Homo sapi
20	443	38.0	918	10	AF061272 Mus muscu
21	441	37.9	753	9	AF411850 Homo sapi
c	439.5	37.7	582	6	AX456965 Sequence
	436	37.4	1037	10	AF387099 Mus muscu
	436	37.4	1418	10	MMU133533 Mus muscu
25	433.5	37.2	615	9	AF200738 Homo sapi
26	430	36.9	938	10	BC006623 Mus muscu
c	427.5	36.7	568	6	AX456969 Sequence
	419.5	36.0	1192	10	BC034893 Mus muscu
29	419	36.0	558	6	AX456968 Sequence
c	400.5	34.4	403	6	AX456970 Sequence
	398.5	34.2	936	9	BC000715 Homo sapi
31	398.5	34.2	968	6	AX083498 Sequence
32	398.5	34.2	997	6	AX375956 Sequence
33	398.5	34.2	2143	9	AB024718 Homo sapi
34	398.5	34.2	1606	10	BC003218 Mus muscu
35	390.5	33.5	2517	10	AB024717 Mus muscu
c	382	32.8	1013	6	AX456966 Sequence
	377.5	32.4	753	6	AX456967 Sequence
38	377.5	32.4	561	6	AX456972 Sequence
39	375.5	32.2	448	6	AX456973 Sequence
40	299.5	25.7	117296	9	AC092865 Homo sapi
41	295.5	25.4	666	6	AX456971 Sequence
42	290.5	24.9	109320	9	AC092746 Homo sapi
43	289.5	24.8	559	6	AX456974 Sequence
44	267.5	23.0	1246	10	MMHL2
45	259	22.2			X53042 Mouse MHL-2

ALIGNMENTS

RESULT 1

AX155225  
LOCUS AX155225 1227 bp DNA linear PAT 22-JUN-2001  
DEFINITION Sequence 3 from Patent WO0136487.  
ACCESSION AX155225  
VERSION AX155225.1 GI:14536703  
KEYWORDS house mouse.  
SOURCE Mus musculus  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
REFERENCE 1 (bases 1 to 1227)  
AUTHORS Schmitz, J., Dzionek, A. and Buck, D.W.  
TITLE Antigen-binding fragments specific for dendritic cells,  
compositions and methods of use thereof antigens recognized thereby  
and cells obtained thereby  
JOURNAL Patent: WO 0136487-A 3 25-MAY-2001;  
Milttenyi Biotec GmbH (DE)  
FEATURES  
source 1..1227  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
146..775  
/note="coding sequence of mouse Dectin-2"  
/codon\_start=1  
/protein\_id="CAC42676.1"  
/db\_xref="GI:14536704"  
/translation="MVQERQSGKGVCTLRLSAAVISMILLSTCFIASCVVYQFI  
MDQSRRLYELHTVHSSLCFSEGTVMSEKMWGCPNHKSGSCYLSTKBNFWS  
SEONCVOMGAHLVINTAEQNFITQOLNESLSYFLGLSDPQGNKQWIDTFFSQN  
VRFWHPHPNLPERCVSIVYWNPSKMGNDVFCDSKHNSICEMKKIYL"  
BASE COUNT 335 a 271 c 276 g 345 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2,97e-73 Length: 1227  
Score: 793.00 Matches: 145  
Percent Similarity: 77.00% Conservative: 19  
Best Local Similarity: 68.08% Mismatches: 41  
Query Match: 68.07% Indels: 8  
DB: 6 Gaps: 4  
US-09-766-511B-53 (1-209) x AX155225 (1-1227)  
QY 1 MetMetGlnGlnGlnProGlnSerThrGluLysArgGly-----TrpLeuSerLeu 18  
Db 146 ATGGTGCAGGAAGACAATCCCAAGG-----AAGGGAGTCTGCTGG---ACCCTG 193  
QY 19 ArgLeuTrpSerValAlaGlyLeuSerIleAlaLeuLeuSerAlaCysPheIleValSer 38  
Db 194 AGACTCTGTCAGCTGCTGTGATTTCCATGTTACTCTTGAGTACTGTTTCATTGCGAGC 253  
QY 39 CysValValThrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHis 58  
Db 254 TGTGTGGTGACTTACCAATTTATATGACAGCCAGCCAGTAGAAGACTATATGAACCTCAC 313  
QY 59 SerTyrHisSerSerLeuThrCysPheSerGluGlyThrLysVal-----ProAlaTrp 76  
Db 314 ACATACCAATTCAGTCTCACCTGCTTCACTGAGTGAAGGACTATGTTGTGAGAAAAATGTGG 373  
QY 77 GlyCysCysProAlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGlu 96  
Db 374 GGATGCTGCCAAATACACTGGAAGTCAATTTGGCTCCAGCTGCTACCTCATTTTCACCAAG 433  
QY 97 GluLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValVal 116  
Db 434 GAGAACTTCTGGAGCACCAGTGACAGAACTGTGTCAGATGGGGGCTCATCTGCTGGTG 493  
QY 117 PheAsnThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyr 136  
Db 494 ATCACTACTGAAGCGGACAGCAATTTATCATCCAGCAGCTGAATGAGTACTTCTTAC 553  
QY 137 PheLeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTnPheAspLysThrPro 156  
Db 137 PheLeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTnPheAspLysThrPro 156

Db 554 TTCTCTGGGTCTTTTCGGATCCACAAAGGTAATGGCAATGGCAATGGATCGATGATCTCT 613  
QY 157 TyrGluLysAsnValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCys 176  
Db 614 TTCAGTCAAAATGTCAGGTCTGGCACCCCATGAACCAATCTTCCAGAGAGCGGTGT 673  
QY 177 AlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThr 196  
Db 674 GTTTCATAGTTTACTTGGGAATCTTCGAAATGGGCTGGAATGATGTTTCTGTGATAGT 733  
QY 197 ArgArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209  
Db 734 AAACAAATTCATATGTGAATGAAGAAGATTACCTA 772  
RESULT 2  
AF240357  
LOCUS AF240357 1227 bp mRNA linear ROD 02-MAY-2000  
DEFINITION Mus musculus strain BALB/c dectin-2 alpha isoform mRNA, complete  
cds.  
ACCESSION AF240357  
VERSION AF240357.1 GI:7677468  
KEYWORDS Mus musculus.  
SOURCE Mus musculus.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
REFERENCE 1 (bases 1 to 1227)  
AUTHORS Ariizumi, K., Shen, G.-L., Shikano, S., Ritter, R. III, Zukas, P.,  
Edelbaum, D., Morita, A. and Takashima, A.  
TITLE Cloning of a second dendritic cell-associated C-type lectin  
(dectin-2) and its alternatively spliced isoforms  
J. Biol. Chem. (2000) In press  
JOURNAL  
REFERENCE 2 (bases 1 to 1227)  
AUTHORS Ariizumi, K. and Takashima, A.  
TITLE Direct Submission  
JOURNAL Submitted (02-MAR-2000) Dermatology, UT Southwestern Medical  
Center, 5323 Harry Hines Blvd, Dallas, TX 75235, USA  
FEATURES  
source 1..1227  
/organism="Mus musculus"  
/strain="BALB/c"  
/db\_xref="taxon:10090"  
/cell\_line="XS52"  
/cell\_type="dendritic cell"  
146..775  
/note="C-type lectin; type II transmembrane protein;  
dendritic cell-associated expression; C-type lectin motif"  
/codon\_start=1  
/product="dectin-2 alpha isoform"  
/protein\_id="AAF67177.1"  
/db\_xref="GI:7677469"  
/translation="MVQERQSGKGVCTLRLSAAVISMILLSTCFIASCVVYQFI  
MDQSRRLYELHTVHSSLCFSEGTVMSEKMWGCPNHKSGSCYLSTKBNFWS  
SEONCVOMGAHLVINTAEQNFITQOLNESLSYFLGLSDPQGNKQWIDTFFSQN  
VRFWHPHPNLPERCVSIVYWNPSKMGNDVFCDSKHNSICEMKKIYL"  
BASE COUNT 335 a 271 c 276 g 345 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2,97e-73 Length: 1227  
Score: 793.00 Matches: 145  
Percent Similarity: 77.00% Conservative: 19  
Best Local Similarity: 68.08% Mismatches: 41  
Query Match: 68.07% Indels: 8  
DB: 10 Gaps: 4  
US-09-766-511B-53 (1-209) x AF240357 (1-1227)  
QY 1 MetMetGlnGlnGlnProGlnSerThrGluLysArgGly-----TrpLeuSerLeu 18  
Db 146 ATGGTGCAGGAAGACAATCCCAAGG-----AAGGGAGTCTGCTGG---ACCCTG 193  
QY 19 ArgLeuTrpSerValAlaGlyLeuSerIleAlaLeuLeuSerAlaCysPheIleValSer 38  
Db 194 AGACTCTGTCAGCTGCTGTGATTTCCATGTTACTCTTGAGTACTGTTTCATTGCGAGC 253  
QY 39 CysValValThrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHis 58  
Db 254 TGTGTGGTGACTTACCAATTTATATGACAGCCAGCCAGTAGAAGACTATATGAACCTCAC 313  
QY 59 SerTyrHisSerSerLeuThrCysPheSerGluGlyThrLysVal-----ProAlaTrp 76  
Db 314 ACATACCAATTCAGTCTCACCTGCTTCACTGAGTGAAGGACTATGTTGTGAGAAAAATGTGG 373  
QY 77 GlyCysCysProAlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGlu 96  
Db 374 GGATGCTGCCAAATACACTGGAAGTCAATTTGGCTCCAGCTGCTACCTCATTTTCACCAAG 433  
QY 97 GluLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValVal 116  
Db 434 GAGAACTTCTGGAGCACCAGTGACAGAACTGTGTCAGATGGGGGCTCATCTGCTGGTG 493  
QY 117 PheAsnThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyr 136  
Db 494 ATCACTACTGAAGCGGACAGCAATTTATCATCCAGCAGCTGAATGAGTACTTCTTAC 553  
QY 137 PheLeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTnPheAspLysThrPro 156  
Db 137 PheLeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTnPheAspLysThrPro 156

```

Db 194 AGACTCTGGTCAGCTGCTGATTTTCATGTTACTTTCAGTACCTGTTTCATTCGGAGC 111
Qy 39 CysValValThrTyrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHis 111
Db 254 TGTGTGGTGACTTACCAATTTATATGGACCACCCAGTAGAAGACTATATGAACCTTAC 111
Qy 59 SerTyrHisSerLeuThrCysPheSerGluGlyThrLysVal-----ProAlaTrp 111
Db 314 ACATACCATTCAGTCTCACCTGCTTCAGTGAAGGAGCTATGTCGACAGAAAAATGTCG 111
Qy 77 GlyCysCysProAlaSerTrpLysSerPheGlySerSerCysTyrPheLeuSerSerGlu 111
Db 374 GGATGCTGCCCAATCACTCGAAGTCACTTGGCTCCAGCTGCTACCTCATCTTCTACCAAG 111
Qy 97 GluLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValVal 111
Db 434 GAGAACTCTGGAGCAGCAGTGGTTCAGTGAAGGAGCTATGTCGAGGGGCTCATCTGGTGG 111
Qy 117 PheAsnThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyr 111
Db 494 ATCAATACTAAGCGGAGCAGAGATTTTCATCACCACGAGCTGAATGAGTCACTTCTTAC 111
Qy 137 PheLeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrPro 111
Db 554 TTCTCTGGGTCTTCGGATCCACAAAGTAATGGCAAAATGGCAATGGATGATGATCTCCT 111
Qy 157 TyrGluLysAsnValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCys 111
Db 614 TTCAGTCAAAATGTGAGTGTCTGGCAGCCCATGACCCCAATCTCCAGAGAGCGGTGT 111
Qy 177 AlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThr 111
Db 674 GTTTCATAGTTTACTGGATCCTTCGAATGGGCTGGGAATGATGTTTCTGTGATAGT 111
Qy 197 ArgArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209
Db 734 AAACACAATTCAATATGTGAATGAAGAAGATTTCACCTA 772

RESULT 3
LOCUS BC023008 1254 bp mRNA linear ROD 07-AUG-2002
DEFINITION Mus musculus, c-type (calcium dependent, carbohydrate recognition domain) lectin, superfamily member 10, clone MGC:35901
IMAGE:5376247, mRNA, complete cds.
ACCESSION BC023008
VERSION BC023008.1 GI:18606437
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1254)
Strausberg, R.
Direct Submission
Submitted (04-FEB-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAP Plate: 53 Row: p Column: 4  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 9910157.

## FEATURES

## Location/Qualifiers

## source

```

1..1254
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N-3"
/clone="MGC:35901 IMAGE:5376247"
/tissue_type="Mammary tumor. MMTV-LTR/INT3 model. 5 month old mouse. Taken by biopsy."
/clone_lib="NCI-CGAP_Mam2"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
146..775
/codon_start=1
/product="c-type (calcium dependent, carbohydrate recognition domain) lectin, superfamily member 10"
/protein_id="AAH23008.1"
/db_xref="GI:18606438"
/db_xref="LocusID:58620"
/translation="MVQEROSQKGVCTLRLNSAAVISMILLSTCFIASCVVYQFI MDQPSRLYLHLVYHSLTCTFSEGTMYSEKMWGCPNHWKSFSGSSCYLIITKDNFWSI SEONCVOMGAHLVYINTEAEQNFITQOLNFSLSYFLGLSDPQNGKQWIDDDTPFSQN VRFWHPHEPNLPPEERCVSIVYWNPSKGNWDFVCDSKHNSICEMKKIYL"

```

## CDS

```

362 a 271 c 276 g 345 c
ORIGIN
Alignment Scores:
Pred. No.: 3,04e-73 Length: 1254
Score: 793.00 Matches: 145
Percent Similarity: 77.00% Conservative: 19
Best Local Similarity: 68.08% Mismatches: 41
Query Match: 68.07% Indels: 8
DB: 10 Gaps: 4

```

US-09-766-511b-53 (1-209) x BC023008 (1-1254)

```

Qy 1 MetMetGlnGluGlnProGlnSerThrGluLysArgGly-----TrpLeuSerLeu 18
Db 146 ATGGTGCAGAAAGACAATCCCAAGG-----AAGGGAGTCTGCTGG---ACCCTG 193
Qy 19 ArgLeuTrpSerValAlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSer 38
Db 194 AGACTCTGGTCAGCTGCTGATTTCCATTTGACTTTCAGTACCTGTTTCATTGCGAGC 253
Qy 39 CysValValThrTyrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHis 58
Db 254 TGTGTGGTGACTTACCAATTTATTTATGGACCCAGCCAGTAGAAGACTATATGAACCTTAC 313
Qy 59 SerTyrHisSerSerLeuThrCysPheSerGluGlyThrLysVal-----ProAlaTrp 76
Db 314 ACATACCATTCAGTCTCACCTGCTTCAGTGAAGGAGCTATGTTGTCAGAAAAATGTCG 373
Qy 77 GlyCysCysProAlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGlu 96
Db 374 GGATGCTGCCCAATCACTCGAAGTCACTTGGCTCCAGCTGCTACCTCATTTCTACCAAG 433
Qy 97 GluLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValVal 116
Db 434 GAGAACTCTGGAGCAGCAGTGGTTCAGTGAAGGAGCTATGTCGAGGGGCTCATCTGGTGG 493
Qy 117 PheAsnThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyr 136
Db 494 ATCAATACTAAGCGGAGCAGAGATTTTCATCACCACGAGCTGAATGAGTCACTTCTTAC 553
Qy 137 PheLeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrPro 156
Db 554 TTCTCTGGGTCTTCGGATCCACAAAGTAATGGCAAAATGGCAATGGATGATGATCTCCT 613

```

```

QY 157 TyrGluLysAsnValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCys 176
   ::::::::::::::::::::
Db 614 TTCAGTCAAAATGTCAGGTTCTGGCACCCCATGAACCCAAATCTCCAGAGAAGCGGTGT 673
   ::::::::::::::::::::

QY 177 AlaserileValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThr 196
   ::::::::::::::::::::
Db 674 GTTCAATAGTTTACTGGAATCCCTCGAAATGGGGCTGGGAATGATGTTTCTGTGATAGT 733
   ::::::::::::::::::::

QY 197 ArgArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209
   ::::::::::::::::::::
Db 734 AACACAATTCAATATGGAATGAAGAAGATTACCTA 772
   ::::::::::::::::::::

RESULT 4
AF240358 1125 bp mRNA linear ROD 02-MAY-2000
LOCUS Mus musculus strain BALB/c dectin-2 beta isoform mRNA, complete
DEFINITION cds, alternatively spliced.
ACCESSION AF240358
VERSION AF240358.1 GI:7677471
KEYWORDS
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1125)
TITL Arizumi,K., Shen,G.-L., Shikano,S., Ritter,R. III, Zukas,P.,
Edelbaum,D., Morita,A. and Takashima,A.
Cloning of a second dendritic cell-associated C-type lectin
(dectin-2) and its alternatively spliced isoforms
J. Biol. Chem. (2000) In press
JOURNAL 2 (bases 1 to 1125)
AUTHORS Arizumi,K. and Takashima,A.
Direct Submission
TITLE Submitted (02-MAR-2000) Dermatology, UT Southwestern Medical
JOURNAL Center, 5323 Harry Hines Blvd, Dallas, TX 75235, USA
FEATURES
source
1..1125
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/cell_line="XS52"
/cell_type="dendritic cell"
146..673
/notes="C-type lectin; type II transmembrane protein;
dendritic cell-associated expression; C-type lectin motif;
alternatively spliced form of dectin-2 alpha mRNA"
/codon_start=1
/product="dectin-2 beta isoform"
/protein_id="AAF67178.1"
/db_xref="GI:7677472"
/translati="MYOEROSQGVGWTLRLWSAAVISMILLSTCFIASCVERKMGCC
CPNHWKSFSGSCYLIETKFNFWSTSEQNCVOMGAHLVWINTAEQNFITQOLNESLSY
FLGLSDPGQNGKWQWIDDPFSQNRVFWHPHEPLPERCVSIYVWPNKSGWNDVFC
DSKHNSICEMKKTYL"
BASE COUNT 305 a 246 c 257 g 317 t
ORIGIN

Alignment Scores:
Pred. No.: 2,61e-59 Length: 1125
Score: 659.00 Matches: 123
Percent Similarity: 65.88% Conservative: 16
Best Local Similarity: 58.29% Mismatches: 34
Query Match: 56.57% Indels: 38
DB: 10 Gaps: 4

US-09-766-511b-53 (1-209) x AF240358 (1-1125)

QY 1 MetMetGlnGluGlnProGlnSerThrGluLysArgGly-----TrpLeuSerLeu 18
   ::::::::::::::::::::
Db 146 ATGTGTCAGGAAGACATCCCAAGGG-----AAGGGAGTCTGCTGG---ACCCTG 193
   ::::::::::::::::::::

QY 19 ArgLeuTrpSerValAlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSer 38
   ::::::::::::::::::::

```

```

Db 194 AGACTCTGCTGACGTGCTCTGTGATTTCCACTGTTCAGTACCTGTTTCATTCCGAGC 253
   ::::::::::::::::::::
QY 39 CysValValThrTyrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHis 58
   ::::::::::::::::::::
Db 254 TGTGTG-----
   ::::::::::::::::::::
QY 59 SerTyrHisSerSerLeuThrCysPheSerGluGlyThrLysValProAlaTrpGlyCys 78
   ::::::::::::::::::::
Db 260 -----GAAAAAATGTGGGATGC 277
   ::::::::::::::::::::
QY 79 CysProAlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGluGluLys 98
   ::::::::::::::::::::
Db 278 TGCCCAATCACTGGAAAGTCATTGGCTCCAGCTCACTTCATCTACCAAGGAGAAC 337
   ::::::::::::::::::::
QY 99 ValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsn 118
   ::::::::::::::::::::
Db 338 TTCTGGAGCACCAGTCAGCAGAACTGTGTTCAGATGGGGCTCATCTGGTGTGATCAAT 397
   ::::::::::::::::::::
QY 119 ThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeu 138
   ::::::::::::::::::::
Db 398 ACTGAAGCGGAGCAGAAATTCATCACCCAGCAGCTGAATGAGTCACTTCTTACTTCCTG 457
   ::::::::::::::::::::
QY 139 GlyLeuSerAspProGlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrProTyrGlu 158
   ::::::::::::::::::::
Db 458 GGTCTTTCGGATCCACAGGTAATGGCAATGGCAATGGATCGATGATCTCTTTCAGT 517
   ::::::::::::::::::::
QY 159 LysAsnValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAlaSer 178
   ::::::::::::::::::::
Db 518 CAAATGTCTCAGTCTCTGGCACCCCATGAACCAATCTCCAGAGCGGTGTGTTC 577
   ::::::::::::::::::::
QY 179 IleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArg 198
   ::::::::::::::::::::
Db 578 ATAGTTTACTGGAATCTTCGAAATGGGCTGGAATGATGTTTCTCTGATAGTAGTAACAC 637
   ::::::::::::::::::::
QY 199 AsnSerIleCysGluMetAsnLysIleTyrLeu 209
   ::::::::::::::::::::
Db 638 AATTCAATATGTGAATGAAGAAGATTACCTA 670
   ::::::::::::::::::::

RESULT 5
AF325459 851 bp mRNA linear PRI 02-DEC-2001
LOCUS Homo sapiens dendritic lectin (CLECSF11) mRNA, complete cds,
DEFINITION alternatively spliced.
ACCESSION AF325459
VERSION AF325459.1 GI:17225336
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE 1 (bases 1 to 851)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Arce,I., Roda-Navarro,P., Montoya,M.C., Hernanz-Falcon,P.,
Puig-Kroger,A. and Fernandez-Ruiz,E.
Molecular and genomic characterization of human DLEC, a novel
member of the C-type lectin receptor gene family preferentially
expressed on monocyte-derived dendritic cells
Eur. J. Immunol. 31 (9), 2733-2740 (2001)
JOURNAL 2 (bases 1 to 851)
MEDLINE 21426806
PUBMED 11536172
REFERENCE 2 (bases 1 to 851)
AUTHORS Arce,I., Hernanz,P. and Fernandez-Ruiz,E.
Direct Submission
TITLE Submitted (01-DEC-2000) Biologia Molecular, Hospital de la
JOURNAL Princesa, Diego de Leon, 62, Madrid 28006, Spain
FEATURES
Location/Qualifiers
1..851
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12"
/map="12p12.3-p13.2"
1..851
/gene="CLECSF11"
27..668
CDS

```



```
/genes="CLECSF11"
/notes="C-type lectin receptor; DLEC; alternatively
spliced"
/codon_start=1
/product="dendritic lectin"
/protein_id="AAL37358.1"
/db_xref="GI:17225337"
/translacion="MVPEEPQDREKGLWFLQKVMASVILLSCYFVSSVVP
NFMYSKTVKRLSKLREYQYHPSLFCVMEKDIKDWSCPTPTWFSQSCYFISFGMQ
SWTKSQKNCVMGADLVINTREODFIIONLKRNSSVFLGLSDPGGRHHQWVDOT
YNENVTWHSWGPNLDERCALINERSSEEWGNDIHCVRPQKSICKMKKIYI"
BASE COUNT 242 a 169 c 192 g 246 t
ORIGIN
```

```
Alignment Scores:
Pred. No.: 3.7e-57 Length: 851
Score: 637.00 Matches: 115
Percent Similarity: 68.08% Conservative: 30
Best Local Similarity: 53.99% Mismatches: 64
Query Match: 54.68% Indels: 4
DB: 9 Gaps: 2
```

US-09-766-511B-53 (1-209) x AF325459 (1-851)

```
Qy 1 MetMetGlnGlnGlnProGlnSerThrGluLys---ArgGlyTrpLeuSerLeuArg 19
Db 27 ATGGTGCCCTCAAGAAGAGCCTCAAGACCGAGAGAAAGGACTCTGGTGGTCCAGTTGAAG 86
Qy 20 LeuTrpSerValAlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSerCys 39
Db 87 GTCTGGTCCATGGCAGTCGTCATCCATCTTGTCTCAGTCTGTCTGTTTCTACTGTGATTCT 146
Qy 40 ValValThrThrHisPheThrGlyGluThrGlyLysArgLeuSerGluLeu----- 57
Db 147 GTGGTGCCCTCAACAATTTATGTATACAAAACTGTCAAGAGGCTGTCCAAAGTTACAGAG 206
Qy 58 ---HisSerTyrHisSerLeuThrCysPheSerGluGlyThrLysValProAlaTrp 76
Db 207 TATCAACAGTATCATCAAGCCTGACCTGCGTCATGGAAGAAAGGACATAGAAGATTGG 266
Qy 77 GlyCysProAlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGlu 96
Db 267 AGCTGTGCCCCAACCCCTTGGACTTCATTTACGTCTAGTTGTCTACTTTATTCTACTGGG 326
Qy 97 GluLysValTrpSerLysSerGluGlnAsnCysValGluMetClyAlaHisLeuValVal 116
Db 327 ATCAATCTTGGACTAAGAGTCAAAAGAACTGTTCTGTGATGGGGCTGATCTGGTGGTG 386
Qy 117 PheAsnThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyr 136
Db 387 ATCAACACCGAGGAAGACAGGATTTCATCATTCAGAAATCTGAAAGAAATCTTCTTAT 446
Qy 137 PheLeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrPro 156
Db 447 TTTCTGGGGCTGTGCATGTCAGGGGTCGCGGACATTTGGCAATGGGTGTGACACACCA 506
Qy 157 TyrGluLysAsnValArgPheThrHisLeuGlyGluProAsnHisSerAlaGluGlnCys 176
Db 507 TACAATAAATGTGCATCTCGGACTCAGGTGAACCAATTAACCTTGATGAGCGTTGT 566
Qy 177 AlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThr 196
Db 567 GCGATAATAAATTTCCGTTCTTCCAGAAGAAATGGGGCTGGAATGACATTCATCTCATGTA 626
Qy 197 ArgArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209
Db 627 CCTCAGAAGTCAATTTTGAAGATGAAGAGATCTACATA 665
```

RESULT 6

```
AX155223 LOCUS 1312 bp mRNA linear PAT 22-JUN-2001
DEFINITION Sequence 1 from Patent WO0136487.
ACCESSION AX155223
```

```
VERSION AX155223.1 GI:14536702
KEYWORDS human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE Schmitz,J., Dzionek,A. and Buck,D.W.
Antigen-binding fragments specific for dendritic cells,
compositions and methods of use thereof antigens recognized thereby
and cells obtained thereby
JOURNAL Patent: WO 0136487-A 1 25-MAY-2001;
FEATURES Miltenyi Biotec GmbH (DE)
SOURCE Location/Qualifiers
1. 1312
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="BDCA-2 cDNA sequence"
BASE COUNT 374 a 301 c 297 g 340 t
ORIGIN
Alignment Scores:
Pred. No.: 6.22e-57 Length: 1312
Score: 637.00 Matches: 115
Percent Similarity: 68.08% Conservative: 30
Best Local Similarity: 53.99% Mismatches: 64
Query Match: 54.68% Indels: 4
DB: 6 Gaps: 2
US-09-766-511B-53 (1-209) x AX155223 (1-1312)
Qy 1 MetMetGlnGlnGlnProGlnSerThrGluLys---ArgGlyTrpLeuSerLeuArg 19
Db 491 ATGGTGCCCTGAAGAAGACGCTCAAGACCGAGAGAAAGGACTCTGGTGGTCCAGTTGAAG 550
Qy 20 LeuTrpSerValAlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSerCys 39
Db 551 GTCGTGCTCATGGCAGTCGTCATCCATCTTGTCTCAGTCTGTCTTCTACTGTGAGTTCT 610
Qy 40 ValValThrThrHisPheThrGlyGluThrGlyLysArgLeuSerGluLeu----- 57
Db 611 GTGGTGCCCTCAACAATTTATGTATACAAAACTGTCAAGAGGCTGTCCAAAGTTACAGAG 670
Qy 58 ---HisSerTyrHisSerLeuThrCysPheSerGluGlyThrLysValProAlaTrp 76
Db 671 TATCAACAGTATCATCAAGCCTGACCTGCGTCATGGAAGAAAGGACATAGAAGATTGG 730
Qy 77 GlyCysProAlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGlu 96
Db 731 AGCTGTGCCCCAACCCCTTGGACTTCATTTACGTCTAGTTGTCTACTTTATTCTACTGGG 790
Qy 97 GluLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValVal 116
Db 791 ATGCAATCTTGGACTAAGAGTCAAAAGAACTGTTCTGTGATGGGGCTGATCTGGTGGTG 850
Qy 117 PheAsnThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyr 136
Db 851 ATCAACACCGAGGAAGACAGGATTTCATCATTCAGAAATCTGAAAGAAATCTTCTTAT 910
Qy 137 PheLeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrPro 156
Db 911 TTTCTGGGGCTGTGCATGTCAGATCCAGGGGTCGCGGACATTTGGCAATGGGTGTGACACACCA 970
Qy 157 TyrGluLysAsnValArgPheThrHisLeuGlyGluProAsnHisSerAlaGluGlnCys 176
Db 971 TACAATAAATGTGCATCTTGGCACTCAGGTGAACCAATTAACCTTGATGAGCGTTGT 1030
Qy 177 AlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThr 196
Db 1031 GCGATAAATAATTTCCGTTCTTCCAGAAGAAATGGGGCTGGAATGACATTCATCTCATGTA 1090
Qy 197 ArgArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209
::: |||||||::: || |||||||:::
```

```

Db 1091 CCTCAGAAGTCAATTGCAAGATGAAGAAGATCTACATA 1129
RESULT 7
AF293615 1313 bp mRNA linear PRI 14-FEB-2002
LOCUS Homo sapiens blood dendritic cell antigen 2 protein (BDCA2) mRNA,
DEFINITION complete cds.
ACCESSION AF293615
VERSION AF293615.1 GI:17224597
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1313)
AUTHORS Dzionek,A., Sohma,Y., Nagafune,J., Cella,M., Colonna,M.,
Faccetti,F., Gunther,G., Johnston,I., Lanzavecchia,A.,
Nagasaka,T., Okada,T., Vermi,W., Winkels,G., Yamamoto,T., Zysk,M.,
Yamaguchi,Y. and Schmitz,J.
TITLE BDCA-2, a novel plasmacytoid dendritic cell-specific type II C-type
lectin, mediates antigen capture and is a potent inhibitor of
interferon alpha/beta induction
J. Exp. Med. 194 (12), 1823-1834 (2001)
JOURNAL 21615135
MEDLINE 11748283
PUBMED
AUTHORS 2 (bases 1 to 1313)
Sohma,Y., Johnston,I., Dzionek,A., Nagasaka,T., Nagafune,J.,
Okada,T., Hirano,T., Schmitz,J. and Yamaguchi,Y.
TITLE Cloning and characterization of blood dendritic cell antigen 2
(BDCA-2), a C-type lectin expressed on plasmacytoid dendritic cells
JOURNAL Unpublished
AUTHORS 3 (bases 1 to 1313)
Sohma,Y., Johnston,I., Dzionek,A., Nagasaka,T., Nagafune,J.,
Okada,T., Hirano,T., Schmitz,J. and Yamaguchi,Y.
TITLE Direct Submission
JOURNAL Submitted (04-AUG-2000) Miltenyi Biotec GmbH, Friedrich-Ebert-Str.
AUTHORS 68, Bergisch Gladbach 51429, Germany
FEATURES
source
1. .1313
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="plasmacytoid dendritic cell; DC-2"
/gene="BDCA2"
492. .1133
/gene="BDCA2"
/note="C-type lectin"
/codon_start=1
/product="blood dendritic cell antigen 2 protein"
/protein_id="AAL37036.1"
/db_xref="GI:17224598"
/translation="MVPPEEPQDREKGLWMFQKVMASVSVLLLSVCFVTSVVYPH
NFMYSKTVKRLSKLREYQYHPSLTCVMEKDIEDWCCPTWTSFQSCYFSTGCMQ
SWTKSKNCVSMGADLVVINTREODFLONLKNSSYFLGLSDPGRRHWQWVDTP
YNEVTFWHSGEPNLDRCALINFRSEEWGWNIDHCHVPQKSIKMKRIY"
BASE COUNT 374 a 302 c 297 g 340 t
ORIGIN
Alignment Scores:
Pred. No.: 6,22e-57 Length: 1313
Score: 637.00 Matches: 115
Percent Similarity: 68.08% Conservative: 30
Best Local Similarity: 53.99% Mismatches: 64
Query Match: 54.68% Indels: 4
DB: 9 Gaps: 2
US-09-766-511b-53 (1-209) x AF293615 (1-1313)
QY 1 MetMetGlnGlnGlnProGlnSerThrGluLys---ArgGlyTrpLeuSerLeuArg 19
||||| ||||||| ||||||| |||||||
Db 492 ATGGTGCTGAAGAGAGCCCTCAAGCCGACAGAGAAGGACTCTGGTGTCCAGTTGAAG 551
US-09-766-511b-53 (1-209) x AX357481 (1-827)
QY 5 GlnGlnProGlnSerThrGluLys---ArgGlyTrpLeuSerLeuArgLeuTrpSerVal 23

```

```

Db 552 GTCTGTCCATGGCAGCTCGTATCCATCTTGCCTCCTCAGTGTCTGTTCCTACTGTGAGTTCT 611
QY 40 ValValThrTyrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeu----- 57
||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 612 GTGGTCCCTCACAAATTTATGTATACAAAAGTGTCAAGAGGCTGCCAAGTTACGAGAG 671
QY 58 ---HisSerTyrHisSerSerLeuThrCysPheSerGluGlyThrLysValProAlaIaTrp 76
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 672 TATCAACAGTATCATCAAGCCCTGACCTCGCTCATGGAAGAAAGACATGAAGATTGG 731
QY 77 GlyCysCysProAlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGlu 96
||||| ||||||| ||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 732 AGCTGTGTCGCCAACCCCTTGGACTTCATTCAGTCTAGTGTCTACTTTATTTCTACTGG 791
QY 97 GluLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValVal 116
: : : : : ||||||| : : : : : ||||||| : : : : : ||||||| : : : : :
Db 792 ATCAATCTTGGACTAAGAGTCAAAAGAACTGTCTGTGATGGGGCTGATCTGGTGGTG 851
QY 117 PheAsnThrGluAlaGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyr 136
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 852 ATCAACACAGGAGCAACAGGATTTTCATCATTCAGAATCTGAAAAAGAAATTTCTTCTAT 911
QY 137 PheLeuGlyLeuSerAspProGlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrPro 156
||||| ||||||| ||||||| ||| : : : : : ||||||| : : : : : |||||||
Db 912 TTCTCTGGGGCTGTTCAGATCCAGGGGTCGGCGACATTTGGCAATGGGTTGACCAGACCA 971
QY 157 TyrGluLysAsnValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGlnGlnCys 176
||| : : : ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||
Db 972 TACAATGAAAATGTACATCTCGGCACCTCAGGTGAACCAATCACTTCAGTGCAGCGTGT 1031
QY 177 AlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThr 196
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1032 CGCATAATAAATTCGGTCTTCAGAGAAGATGGGCTGGAATGCATTCACATTCATCTCA 1091
QY 197 ArgArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209
: : : ||||||| : : : ||| ||||||| : : :
Db 1092 CCTCAGAAGTCAATTTGCAAGATGAAGAAGATCTACATA 1130
RESULT 8
AX357481 827 bp DNA linear PAT 13-FEB-2002
LOCUS Sequence 19 from Patent WO0172773.
DEFINITION AX357481
ACCESSION AX357481
VERSION AX357481.1 GI:18674538
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Werner,G., Phares,W., Jaritz,M., Lapp,H. and Kalthoff,F.S.
TITLE Organic compounds
JOURNAL Patent: WO 0172773-A 19 04-OCT-2001;
Novartis AG (CH)
FEATURES
Source
Location/Qualifiers
1. .827
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 242 a 164 c 188 g 233 t
ORIGIN
Alignment Scores:
Pred. No.: 1.92e-56 Length: 827
Score: 630.00 Matches: 114
Percent Similarity: 67.94% Conservative: 28
Best Local Similarity: 54.53% Mismatches: 63
Query Match: 54.08% Indels: 4
DB: 6 Gaps: 2
US-09-766-511b-53 (1-209) x AX357481 (1-827)
QY 5 GlnGlnProGlnSerThrGluLys---ArgGlyTrpLeuSerLeuArgLeuTrpSerVal 23

```

```
Db 28 CAAGAGCCTCAAGACCAGAGAAAGGACTCTGGTGGTTCCAGTTGAAGCTCGTCCATG 87
||||:||||| ||||| ||| |||:|||||:|||||:
Qy 24 AlaclyIleSerIleAlaLeuSerAlaCysPheIleValSerCysValThrTyr 43
|||:|||||:|||||:|||||:|||||:|||||:
Db 88 GCAGTCGTATCCATCTTGTCTCCAGTGTCTGTTTCACTGTAGTCTCTGGTGGCTCAC 147
||||:|||||:|||||:|||||:|||||:|||||:
Qy 44 HisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeu-----HisSerTyr 60
||||:|||||:|||||:|||||:|||||:|||||:
Db 148 AATTATGATATAGCAAAATGTCAGAGGCTGTCCAAGTTACGACAGATCAACAGTAT 207
||||:|||||:|||||:|||||:|||||:|||||:
Qy 61 HisSerSerLeuThrCysPheSerGluGlyThrLysValProAlaTrpGlyCysCysPro 80
|||:|||||:|||||:|||||:|||||:|||||:
Db 208 CATCAAGCCTGACCTCGCTCATCGTGAAGGAAGACATAGAAGATTGGAGCTGCTCCCA 267
||||:|||||:|||||:|||||:|||||:|||||:
Qy 81 AlaSerTrpLysSerPheGlySerCysTyrPheIleSerSerGluGluLysValTrp 100
||||:|||||:|||||:|||||:|||||:|||||:
Db 268 ACCCTTGGACTTCATTTCACTAGTGTCTATTTTACTGGGATGCAATCTGG 327
||||:|||||:|||||:|||||:|||||:|||||:
Qy 101 SerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsnThrGlu 120
||||:|||||:|||||:|||||:|||||:|||||:
Db 328 ACTAAGAGTCAAAAGAACTGTTCTGTGATGGGGCTCATCTGGTGTGATCAACACGAG 387
||||:|||||:|||||:|||||:|||||:|||||:
Qy 121 AlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeu 140
||||:|||||:|||||:|||||:|||||:|||||:
Db 388 GAAGAACAGGATTTCATTCATTCAGAAATCTGAAAAAGAAATTTCTTCTATTTTCTGGGGCTG 447
||||:|||||:|||||:|||||:|||||:|||||:
Qy 141 SerAspProGlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsn 160
||||:|||||:|||||:|||||:|||||:|||||:
Db 448 TCAGATCCAGGGGTCGGCGACATTTGGCAATGGGTTTGACACACACATACATGAAAT 507
||||:|||||:|||||:|||||:|||||:|||||:
Qy 161 ValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAlaSerIleVal 180
|||:|||||:|||||:|||||:|||||:|||||:
Db 508 GTACATCTCGGCACCTCAGTGAACCAATACCTTGATGAGCGTTGTCGGATAATAAT 567
||||:|||||:|||||:|||||:|||||:|||||:
Qy 181 PheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArgAsnSer 200
|||:|||||:|||||:|||||:|||||:|||||:
Db 568 TTCGTTCTTCAGAAAGTGGGCTGGAATGACATTCACCTGTCTACCTCAGAAAGTCA 627
||||:|||||:|||||:|||||:|||||:|||||:
Qy 201 IleCysGluMetAsnLysIleTyrLeu 209
||||:|||||:|||||:|||||:|||||:
Db 628 ATTTGCAAGATGAAGAAGATCTACATA 654
||||:|||||:|||||:|||||:|||||:
RESULT 9
AX357483 800 bp DNA linear PAT 13-FEB-2002
LOCUS
DEFINITION Sequence 21 from Patent WO0172773.
ACCESSION AX357483
VERSION AX357483.1 GI:18674539
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE Werner, G., Phares, W., Jaritz, M., Lapp, H. and Kalthoff, F.S.
JOURNAL Organic Compounds
Patent: WO 0172773-A 21 04-OCT-2001;
Novartis AG (CH)
FEATURES
Location/Qualifiers
source 1..800
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 231 a 160 c 183 g 226 t
ORIGIN
Alignment Scores:
Pred. No.: 4.2e-55 Length: 800
Score: 617.00 Matches: 111
Percent Similarity: 67.14% Conservative: 30
Best Local Similarity: 52.86% Mismatches: 63
Query Match: 52.96% Indels: 6
DB: 6 Gaps: 2
```

```
US-09-766-511B-53 (1-209) x AX357483 (1-800)
Qy 6 GlnProGlnSerThrGluLysArg-----GlyTrpLeuSerLeuArgLeuTrpSer 22
||||:||||| |||:|||||:|||||:
Db 2 CCCCCCGGTCGGAAGACCGAGAGAAAGACTCTGGTGGTCCAGTTCCAGGTCCTGGTCC 61
||||:|||||:|||||:|||||:|||||:
Qy 23 ValAlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysValValThr 42
||||:|||||:|||||:|||||:|||||:|||||:
Db 62 ATGGCAGTCGATCATCTTCTCTCAGTGTCTGTTTCACTGTGAGTCTCTGGTGGCT 121
||||:|||||:|||||:|||||:|||||:|||||:
Qy 43 TyrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeu-----HisSer 59
||||:|||||:|||||:|||||:|||||:|||||:
Db 122 CACAATTTATGTATAGCAAAACTCTCAAGAGGCTGTCCAAGTTACGAGAGATATCAACAG 181
||||:|||||:|||||:|||||:|||||:|||||:
Qy 60 TyrHisSerSerLeuThrCysPheSerGluGlyThrLysValProAlaTrpGlyCysCys 79
||||:|||||:|||||:|||||:|||||:|||||:
Db 182 TATATCCAAAGCTTCACTGGCTCATGGAAGGAAGACATAGAAGATTGGAGCTGTGCG 241
||||:|||||:|||||:|||||:|||||:|||||:
Qy 80 ProAlaSerTrpLysSerPheGlySerCysTyrPheIleSerSerGluLysVal 99
||||:|||||:|||||:|||||:|||||:|||||:
Db 242 CCAACCCCTTGCAGTTCATTTCACTAGTGTCTACTTTATTTCTACTGGGATGCAATCT 301
||||:|||||:|||||:|||||:|||||:|||||:
Qy 100 TrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsnThr 119
||||:|||||:|||||:|||||:|||||:|||||:
Db 302 TGGACTAAGAGTCAAAAGAACTGTTCTCTGTATGGGGCTGATCTGGTGTGATCAACACC 361
||||:|||||:|||||:|||||:|||||:|||||:
Qy 120 GluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGly 139
||||:|||||:|||||:|||||:|||||:|||||:
Db 362 AGGGAAGAACAGAGATTTTCATTCATTCAGAAATCTGAAAAAGAAATTTCTTCTATTTCTGGGG 421
||||:|||||:|||||:|||||:|||||:|||||:
Qy 140 LeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLys 159
||||:|||||:|||||:|||||:|||||:|||||:
Db 422 CTGTCAGATCCAGGGGTCGGCGACATTTGGCAATGGGTTGACCAGACACCATCAATGAA 481
||||:|||||:|||||:|||||:|||||:|||||:
Qy 160 AsnValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAlaSerIle 179
||||:|||||:|||||:|||||:|||||:|||||:
Db 482 AATGTACATTTCTGGCAGTCAACCAATAACCTTGATGAGCGTTGTGCGATAATA 541
||||:|||||:|||||:|||||:|||||:|||||:
Qy 180 ValPheTrpLysProThrGlyTrpAsnAspValIleCysGluThrArgArgAsn 199
||||:|||||:|||||:|||||:|||||:|||||:
Db 542 AATTTCCGTTCTTCAGAAAGATGGGGCTGGAATGACATTCATCTGTCATGTACCTCAGAAG 601
||||:|||||:|||||:|||||:|||||:|||||:
Qy 200 SerIleCysGluMetAsnLysIleTyrLeu 209
||||:|||||:|||||:|||||:|||||:|||||:
Db 602 TCAATTGCAAGATGAAGAGATCTACATA 631
||||:|||||:|||||:|||||:|||||:|||||:
RESULT 10
AF240359 1125 bp mRNA linear ROD 02-MAY-2000
LOCUS
DEFINITION AF240359 cds, alternatively spliced.
ACCESSION AF240359
VERSION AF240359.1 GI:7677474
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 1125)
JOURNAL Arizumi, K., Shen, G.-L., Shikano, S., Ritter, R. III, Zukas, P.,
REFERENCE Arizumi, K., Morita, A. and Takashima, A.
AUTHORS Edelbaum, D., Morita, A. and Takashima, A.
TITLE Cloning of a second dendritic cell-associated C-type lectin
JOURNAL (dectin-2) and its alternatively spliced isoforms
FEATURES J. Biol. Chem. (2000) In press
REFERENCE 2 (bases 1 to 1125)
AUTHORS Arizumi, K. and Takashima, A.
TITLE Direct Submission
JOURNAL Submitted (02-MAR-2000) Dermatology, UT Southwestern Medical
Center, 5323 Harry Hines Blvd, Dallas, TX 75235, USA
FEATURES Location/Qualifiers
source 1..1125
/organism="Mus musculus"
```



```

Db 218 CAATCTTGGACTAAGAGTCAAAGAAGCTGTTCTGTGTATGGGGCTGATCTGGTGTATC 277
Qy 118 AsnThrGluAlaGluGlnAsnPhelIeValGlnGlnLeuAsnGluSerPheSerTyrPhe 137
Db 278 AACACCGGGAGACAGGATTCTCATCTCAGAAATCTGAAAGAAATCTTCTTATTTT 337
Qy 138 LeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrProTyr 157
Db 338 CTGGGGCTGTGATGATCGAGGGGTCGGCGACATTGGCAATGGGTGTGACACAGACCATAC 397
Qy 158 GluLysAsnValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAla 177
Db 398 AATGAAATTCATCATCTTGGCATCGAGTGAACCCCAATAACCTTGATGAGCGTGTGGC 457
Qy 178 SerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArg 197
Db 458 ATAATAATTTCCGTTCTTTCAGAAGATGGGGCTGGAAATGACATTCATCTGTCATGTA 517
Qy 198 ArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209
Db 518 CAGAAGTCAATTTGCAAGATGAAGAAGATCTACATA 553

RESULT 12
LOCUS AX456975 AX456975 444 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 14 from Patent WO0232958.
ACCESSION AX456975
VERSION AX456975.1 GI:21715770
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Smolyar, A.
TITLE Polynucleotide and polypeptide sequences of human dendritic cell
immunoreceptors
JOURNAL Patent: WO 0232958-A 14 25-APR-2002;
Bayer Aktiengesellschaft (DE)
FEATURES
source 1..444
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 123 a 86 c 101 g 134 t
ORIGIN
Alignment Scores:
Pred. No.: 5.75e-39 Length: 444
Score: 459.50 Matches: 80
Percent Similarity: 67.57% Conservative: 20
Best Local Similarity: 54.05% Mismatches: 47
Query Match: 39.44% Indels: 1
DB: 6 Gaps: 1

US-09-766-511B-53 (1-209) x AX456975 (1-444)
Qy 63 SerLeuThrCysPheSerGluGlyThrLysVal---ProAlaTrpGlyCysCysProAla 81
Db 1 TCTGTTAATATTTTCTTCTTCTATATCTATTGGGAAGATTGGAGCTGCTCCCAACC 60
Qy 82 SerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGluGluLysValTrpSer 101
Db 61 CCTTGGACTTCATTCATCTAGTCTAGTCTACTTTTCTACTTTTCTACTGGGATGCAATCTGGACT 120
Qy 102 LysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsnThrGluAla 121
Db 121 AAGAGTCAAAAGAACTGTTCTGTATGGGGCTGATCTGGTGTGATCAACACAGGGAA 180
Qy 122 GluGlnAsnPhelIeValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeuSer 141
Db 181 GAACAGGATTTCATCATCTCAGAAATCTTCTTATTTTCTGGGGCTGTCA 240
Qy 142 AspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsnVal 161

```

```

Db 241 GATCCAGGGGTCGGGACATTGGCAATGGTGTGACACACACCATACATCAATAATGTC 300
Qy 162 ArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAlaSerIleValPhe 181
Db 301 ACATTCTGGCACTCAGGTGAACCAATAACCTTGATGAGCGTGTGCGATAATAATTTTC 360
Qy 182 TrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArgAsnSerIle 201
Db 361 CGTTCTTCAGAAGATGGGGCTGGAAATGACATTCCTGTCATGTCACCTCAGAAGTCAATT 420
Qy 202 CysGluMetAsnLysIleTyrLeu 209
Db 421 TGCAAGATGAAGAAGATCTACATA 444

RESULT 13
LOCUS AX456962 AX456962 402 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 1 from Patent WO0232958.
ACCESSION AX456962
VERSION AX456962.1 GI:21715759
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Smolyar, A.
TITLE Polynucleotide and polypeptide sequences of human dendritic cell
immunoreceptors
JOURNAL Patent: WO 0232958-A 1 25-APR-2002;
Bayer Aktiengesellschaft (DE)
FEATURES
source 1..402
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 114 a 80 c 96 g 112 t
ORIGIN
Alignment Scores:
Pred. No.: 5.76e-39 Length: 402
Score: 459.00 Matches: 77
Percent Similarity: 70.90% Conservative: 18
Best Local Similarity: 57.46% Mismatches: 39
Query Match: 39.40% Indels: 0
DB: 6 Gaps: 0

US-09-766-511B-53 (1-209) x AX456962 (1-402)
Qy 76 TrpGlyCysCysProAlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSer 95
Db 1 TGGAGCTGCTGCCAACCCCTTGGACTTCATTTTCAGTCTAGTGTGCTACTTTATTTCTACT 60
Qy 96 GluGluLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuVal 115
Db 61 GGGATGCAATCTGGGACTAAGAGTCAAAAGAACTGTTCTGTGATGGGGCTCATCTGGTG 120
Qy 116 ValPheAsnThrGluAlaGluGlnAsnPhelIeValGlnGlnLeuAsnGluSerPheSer 135
Db 121 GTGATCAACACCGGAAGAACAGGATTTCATCTCAGAATCTGAAAAGAAATTTCTTCT 180
Qy 136 TyrPheLeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThr 155
Db 181 TATTTTCTGGGCTGTGATCCAGGGGTGCGGCACATTTGGCATGGGTGACACGACA 240
Qy 156 ProTyrGluLysAsnValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGln 175
Db 241 CCATCAATGAAATGTACATTTCTGGCACTCAGGTGAACCCCAATAACCTTGATGAGCGT 300
Qy 176 CysAlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGlu 195
Db 301 TGTGGCATATAAATTTCCGTTCTTTCAGAAGAAATGGGGCTGGAAATGACATTCATCTGTCAT 360

```

Qy	196	ThrArgArgAsnSerIleCysGluMetAsnLysIleTyrLeu	209
Db	361	GTACCTCAGAAGTCATTTGCAAGATGAAGAATCTACATA	402
RESULT 14			
AF109146		linear	PRI 24-OCT-2000
LOCUS		mRNA	
DEFINITION	Homo sapiens C-type lectin superfamily 6 (CLECSF6)	mRNA, complete cds.	
ACCESSION	AF109146		
VERSION	AF109146.1	GI:6502534	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS	Richard,M. and Beaulieu,A.D.		
TITLE	A novel C-type lectin expressed in GM-CSF stimulated neutrophils		
JOURNAL	Unpublished		
REFERENCE	1. (bases 1 to 1133)		
AUTHORS	Richard,M. and Beaulieu,A.D.		
TITLE	Direct Submission		
JOURNAL	Submitted (24-NOV-1998) Dept. of Medicine, Centre Hospitalier de l'Universite Laval, 2705 Boul. Laurier, Sainte-Foy, Quebec G1V 4G2, Canada		
FEATURES			
source			
gene			
CDS			
BASE COUNT	332 a	214 c	239 g
ORIGIN			
Alignment Scores:			
Pred. No.:	2,54e-38	Length:	1133
Score:	458.00	Matches:	83
Percent Similarity:	62.57%	Conservative:	34
Best Local Similarity:	44.39%	Mismatches:	66
Query Match:	39.31%	Indels:	4
DB:	9	Gaps:	3
US-09-766-511B-53 (1-209) x AF109146 (1-1133)			
Qy	26	IleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysValValThrTyrHisPhe	45
Db	248	ATATTTTCTCGTATTGGCAATCTCATTTTATTGCTTTTGCTATTTCTTTCAA---	304
Qy	46	ThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHisSerTyrHisSerSerLeuThr	65
Db	305	AAATATTCTAGCTTCTTGAATAAAGACTACAAAAGAGCTGTTTCATACACATTGGAG	364
Qy	66	CysPheSerGlyGluThrLysVal-----ProAlaTrpGlyCysCysProAlaSerTrp	83
Db	365	TGTGTGAAAAAATAATGCCCGTGAAGACACACCTGGAGCTGTGCCAACGAATTCG	424
Qy	84	LysSerPheGlySerSerCysTyrPheIleSerSerGluLysValTrpSerLysSer	103
Db	425	AAGTCATTAGTCCACTGCTACTTATTTACTTACTGCATCTTGGCAAGACAGT	484







GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 20, 2003, 19:29:57 : Search time 58 Seconds  
(without alignments)  
1835.321 Million cell updates/sec

Title: US-09-766-511B-53  
Perfect score: 1165  
Sequence: 1 MMQEQOPSTKRGWLSRL.....NDVICETRNICEWNKIYL 209

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 424239 segs, 254661826 residues

Total number of hits satisfying chosen parameters: 848478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_p2n.model -DEV=xlh  
-DB=/cgn2\_1/USPTO.spool/US09766511/runat\_19022003\_094339\_29351/app\_query.fasta\_1.391  
-DB-PublishedApplications\_NA -QPMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cgi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09766511@cgn\_1\_1\_28@runat\_19022003\_094339\_29351  
-NCPU=6 -ICPU=3 -NO\_XLPXY -NO\_MMAPP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications\_NA:

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	765.5	65.7	693	10 US-09-833-381-994	Sequence 994, Appl
C 2	653.5	56.1	528	10 US-09-833-381-995	Sequence 995, Appl
3	637	54.7	642	12 US-10-090-466-1	Sequence 1, Appl
4	535.5	46.0	549	12 US-10-090-466-3	Sequence 3, Appl

5	458	39.3	1091	10 US-09-764-870-199	Sequence 199, App
6	458	39.3	1096	10 US-09-764-870-38	Sequence 38, Appl
7	458	39.3	1104	9 US-09-862-802-1	Sequence 1, Appl
8	457.5	39.3	1036	10 US-09-833-381-119	Sequence 119, App
9	440.5	37.8	758	10 US-09-833-381-1339	Sequence 1339, App
c	436	37.4	1418	9 US-09-862-802-7	Sequence 7, Appl
10	398.5	34.2	968	9 US-09-965-529-40	Sequence 40, Appl
11	398.5	34.2	997	9 US-09-905-291A-376	Sequence 376, App
12	398.5	34.2	997	9 US-09-902-853-376	Sequence 376, App
13	398.5	34.2	997	9 US-09-907-824-376	Sequence 376, App
14	398.5	34.2	997	9 US-09-907-841-376	Sequence 376, App
15	398.5	34.2	997	9 US-09-904-011-376	Sequence 376, App
16	398.5	34.2	997	9 US-10-174-590-23	Sequence 23, Appl
17	398.5	34.2	997	9 US-10-176-758-23	Sequence 23, Appl
18	398.5	34.2	997	9 US-10-175-737-23	Sequence 23, Appl
19	398.5	34.2	997	9 US-09-906-742-376	Sequence 376, App
20	398.5	34.2	997	9 US-10-173-706-23	Sequence 23, Appl
21	398.5	34.2	997	9 US-10-175-738-23	Sequence 23, Appl
22	398.5	34.2	997	9 US-10-175-752-23	Sequence 23, Appl
23	398.5	34.2	997	9 US-10-176-482-23	Sequence 23, Appl
24	398.5	34.2	997	9 US-10-176-757-23	Sequence 23, Appl
25	398.5	34.2	997	9 US-10-176-913-23	Sequence 23, Appl
26	398.5	34.2	997	9 US-10-180-552-23	Sequence 23, Appl
27	398.5	34.2	997	9 US-10-180-557-23	Sequence 23, Appl
28	398.5	34.2	997	9 US-09-906-838-376	Sequence 376, App
29	398.5	34.2	997	9 US-09-907-613-376	Sequence 376, App
30	398.5	34.2	997	9 US-09-907-942-376	Sequence 376, App
31	398.5	34.2	997	9 US-10-173-700-23	Sequence 23, Appl
32	398.5	34.2	997	9 US-10-174-572-23	Sequence 23, Appl
33	398.5	34.2	997	9 US-10-174-579-23	Sequence 23, Appl
34	398.5	34.2	997	9 US-10-174-582-23	Sequence 23, Appl
35	398.5	34.2	997	9 US-10-174-588-23	Sequence 23, Appl
36	398.5	34.2	997	9 US-10-175-739-23	Sequence 23, Appl
37	398.5	34.2	997	9 US-10-175-740-23	Sequence 23, Appl
38	398.5	34.2	997	9 US-10-175-743-23	Sequence 23, Appl
39	398.5	34.2	997	9 US-10-176-488-23	Sequence 23, Appl
40	398.5	34.2	997	9 US-10-176-492-23	Sequence 23, Appl
41	398.5	34.2	997	9 US-10-176-747-23	Sequence 23, Appl
42	398.5	34.2	997	9 US-10-176-750-23	Sequence 23, Appl
43	398.5	34.2	997	9 US-10-176-985-23	Sequence 23, Appl
44	398.5	34.2	997	9 US-10-176-987-23	Sequence 23, Appl
45	398.5	34.2	997	9 US-10-176-987-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1  
US-09-833-381-994/c  
; Sequence 994, Application US/09833381  
; Patent No. US20020132090A1  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; CURRENT FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 994  
; LENGTH: 693  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(693)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-833-381-994

Alignment Scores:  
Pred. No.: 5e-78 Length: 693  
Score: 765.50 Matches: 138  
Percent Similarity: 78.28% Conservative: 17

Best Local Similarity: 69.70% Mismatches: 39  
Query Match: 65.71% Indels: 4  
DB: 10 Gaps: 2

US-09-766-511b-53 (1-209) x US-09-833-381-994 (1-693)

```
QY 15 TrpLeuSerLeuArgLeuTrpSerValAlaGlyIleSerIleAlaLeuLeuSerAlaCys 34
   ||| :|||:|||||:|||||: ||| :|||:|||||: ||| :|||:|||||: |||
Db 674 TGG--ACCTGAGACTCTGGTCAGCTGCTGTGATTTCCATGTTACTTCTTGAGTACCTGT 618
QY 35 PheIleValSerCysValValThrTyrHisPheThrTyrGlyGluThrGlyLysArgLeu 54
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 617 TTCATTGCCAGCTGTGTGTGACTTACCAATTTATATGGACCACCCAGGACTATGGTGCA 558
QY 55 SerGluLeuHisSerTyrHisSerSerLeuThrCysPheSerGluGlyThrLysVal--- 73
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 557 TATGAATTCACACATACCATTCCAGTCTCACCTGCTTCAGTGAAGGACTATGGTGCA 498
QY 74 ---ProAlaTrpGlyCysProAlaSerTrpLysSerPheGlySerSerCysTyrPhe 92
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 497 GAAAAAATGTGGGATGCTGCCCAATCACCTGGAAGTCATTTGGCTCCAGCTGTACCTC 438
QY 93 IleSerSerGluGlyLysVal-TripSerLysSerGluGlnAsnCysValGluMetGlyAl 112
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 437 ATTCTACCAAGAGAACTCTGGAGCACCACTGAGCAGAACTGTTCATATGGGGGC 378
QY 112 aHisLeuValValPheAsnThrGluAlaGluGlnAsnPheIleValGlnLeuAsnG1 132
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 377 TCATCTGGTGGTGATCAATACTGAAGCGAGCAGAGAATTTCAACCCAGCAGCTGAATGA 318
QY 132 uSerPheSerTyrPheLeuGlyLeuSerAspProGlnGlyAsnAsnAsnTrpGlnTrpI1 152
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 317 GTCACTTTNTACTTCCTGGGTCCTTCGGATCCACAAAGGTAATGGCAAATGGCAT 258
QY 152 eAspLysThrProTyrGluLysAsnValArgPheThrHisLeuGlyGluProAsnHisSe 172
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 257 CGATGATACCTCTTCAGTCAAAATGTCAGGTCTGGCACCCCATGAACCAATCTTCC 198
QY 172 rAlaGluGlnCysAlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspVa 192
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 197 AGAAGAGCGGTGTGTTCATAGTTTACTTGGAAATCCTTCGAAATGGGGCTGGAATGATG 138
QY 192 lIleCysGluThrArgArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 137 TTTCTGTGATGAACACAAATTCATATGTCAAATGAAGAAGATTTACCTA 86
```

## RESULT 2

```
; US-09-833-381-995/C
; Sequence 995, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: NO. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 995
; LENGTH: 528
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(528)
; OTHER INFORMATION: n = A,T,C or G
```

## US-09-833-381-995

Alignment Scores:  
Pred. No.: 1,78e-65 Length: 528  
Score: 653.50 Matches: 120

Percent Similarity: 78.11% Conservative: 12  
Best Local Similarity: 71.01% Mismatches: 34  
Query Match: 56.09% Indels: 3  
DB: 10 Gaps: 2

US-09-766-511b-53 (1-209) x US-09-833-381-995 (1-528)

```
QY 15 TrpLeuSerLeuArgLeuTrpSerValAlaGlyIleSerIleAlaLeuLeuSerAlaCys 34
   ||| :|||:|||||:|||||: ||| :|||:|||||: ||| :|||:|||||: |||
Db 509 TGG--ACCTGAGACTCTGTGTCAGCTGCTGTGATTTCCATGTTACTTCTTGAGTACCTGT 453
QY 35 PheIleValSerCysValValThrTyrHisPheThrTyrGlyGluThrGlyLysArgLeu 54
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 452 TTCATTCCGAGCTGTGTGTGACTTACCAATTTATATGGACCACCCAGTGAAGACTA 393
QY 55 SerGluLeuHisSerTyrHisSerSerLeuThrCysPheSerGluGlyThrLysVal--- 73
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 392 TATGAATTCACACATACCATTCCAGTCTCACCTGCTTCAGTGAAGGACTATGGTGCA 333
QY 74 ---ProAlaTrpGlyCysCysProAlaSerTrpLysSerPheGlySerSerCysTyrPhe 92
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 332 GAAAAAATGTGGGATGCTGCCCAATCACCTGGAAGTCATTTGGCTCCAGCTGTACCTC 273
QY 93 IleSerSerGluGlyLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAla 112
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 272 ATTCTACCAAGGAGAACTTCTGGAGCACCACTGAGCAAGTCATTTGGCTCCAGCTGTACCTC 213
QY 113 HisLeuValValPheAsnThrGluAlaGluGlnAsnPheIleValGlnLeuAsnGlu 132
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 212 CATCTGGTGGTGATCAATACTGAAGCGGAGCAGAAATTTCAACCCAGCAGCTGAATGAG 153
QY 133 SerPheSerTyrPheLeuGlyLeuSerAspProGlnGlyAsnAsnAsnTrpGlnTrpIle 152
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 152 TCATCTTCTTACTCTCTGGGTCCTTCGGATCCACAAGGTAATGGCAAATGGCAATGATC 93
QY 153 AspLysThrProTyrGluLysAsnValArgPheThrHisLeuGlyGluProAsnHisSer 172
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 92 GATGATACCTCTTTCAGTCAAAATGTCAGGTCTGGCACCCCATGAACCAATCTTCCA 33
QY 173 AlaGluGlnCysAlaSerIleValPhe 181
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 32 GAAGACGGGTGCTGTTCAATAGATTTCAC 6
```

## RESULT 3

```
; US-10-090-466-1
; Sequence 1, Application US/10090466
; Patent No. US20020137914A1
; GENERAL INFORMATION:
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Mathur, Brian
; APPLICANT: Cullinan, Emily B.
; TITLE OF INVENTION: NO. US20020137914A1el Human Dectin Proteins and Polynucleotide
; FILE REFERENCE: LEX-0315-USA
; CURRENT APPLICATION NUMBER: US/10/090,466
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 60/274,961
; PRIOR FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 642
; TYPE: DNA
; ORGANISM: homo sapiens
; US-10-090-466-1
```

## Alignment Scores:

Pred. No.: 1,79e-63 Length: 642  
Score: 637.00 Matches: 115  
Percent Similarity: 68.08% Conservative: 30  
Best Local Similarity: 53.99% Mismatches: 64  
Query Match: 54.68% Indels: 4  
DB: 12 Gaps: 2

US-09-766-511B-53 (1-209) x US-10-090-466-1 (1-642)

```
QY 1 MetMetGlnGluGlnProGlnSerThrGluLys--ArgGlyThrLeuSerLeuArg 19
||||: |||:||||| ||||| ||||| |||||
Db 1 ATGGTGCTGAAGAAGAGCCTCAAGACCGAGAGAAAGGACTCTGGTGGTTCAGAGTTGAAG 60

QY 20 LeuTrpSerValAlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSerCys 39
:|||||:||||| |||:||||| ||||| ||||| |||||
Db 61 GTCTGGTCCATGCGAGTCGATCCATCTTGCCTCAGTGTCTGTTTCACTGTGAGTTCT 120

QY 40 ValValThrTyrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeu----- 57
|||||:||||| |||:||||| ||||| ||||| |||||
Db 121 GTGGTGCTTCACAAATTTTATGTATCAACAACTGTCAAGAGGCTGTCCAAGTTACGAGAG 180

QY 58 ---HisSerTyrHisSerLeuThrCysPheSerGluGlyThrLysValProAlaTrp 76
||||| ||||| ||||| ||||| ||||| |||||
Db 181 TATCAACAGTATCATCAAGCCTGACCTCGGTCAATGGAAGGAAGACATAGAAGATTGG 240

QY 77 GlyCysCysProAlaSerTrpLysSerPheGlySerCysTyrPheIleSerSerGlu 96
||||| ||||| ||||| ||||| ||||| |||||
Db 241 AGCTGTGCTGCCAACCCCTTGGACTTCATTTTCAGTCTAGTTGCTACTTTATTCTACTGGG 300

QY 97 GluLysValTrpSerLysSerGluGlnAsnCysValGluMetClyAlaHisLeuValVal 116
||||| ||||| ||||| ||||| ||||| |||||
Db 301 ATGCAATCTGGAGCTAAGAGTCAAAAGAACTGTTCTGTCTGGGGCTGATCTGTGGTG 360

QY 117 PheAsnThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyr 136
||||| ||||| ||||| ||||| ||||| |||||
Db 361 ATCAACACCGAGGAAGACAGGATTTTCATCTCAGAACTCAAAAGAAATTTCTTCTTAT 420

QY 137 PheLeuGlyLeuSerAspProGlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrPro 156
||||| ||||| ||||| ||||| ||||| |||||
Db 421 TTCTCTGGGGCTGTCTAGATCCAGGGCTCGCGGACATTTGGCAATGGGTTTCAACACACCA 480

QY 157 TyrGluLysAsnValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCys 176
||||| ||||| ||||| ||||| ||||| |||||
Db 481 TACAATGAAATGTACATCTTGGCAGTCAAGTGAACCCCAATACCTTGATGAGCGTTGT 540

QY 177 AlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThr 196
|||| ||| ||| ||||| ||||| ||||| |||||
Db 541 GCGATAATAATTTCCGTTCTTCAGAGAATGGGGCTGGAATGACATTCACCTGTCATGTA 600

QY 197 ArgArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209
:|: |||||:||||| ||||| |||||
Db 601 CCTCAGAAGTCAATTTGCAAGATGAAGAAGATCTACATA 639
```

## RESULT 4

```
; Sequence 3, Application US/10090466
; Patent No. US20020137914A1
; GENERAL INFORMATION:
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Mather, Brian
; APPLICANT: Cullinan, Emily B.
; TITLE OF INVENTION: No. US20020137914A1el Human Dectin Proteins and Polynucleotides B
; FILE REFERENCE: LEX-0315-USA
; CURRENT APPLICATION NUMBER: US/10/090,466
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 60/274,961
; PRIOR FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 549
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-090-466-3
```

Alignment Scores: 4.86e-52 Length: 549  
Pred. No.: 535.50 Matches: 94  
Score: 68.02% Conservative: 23  
Percent Similarity: 54.65% Mismatches: 52

Query Match: 45.97% Indels: 3  
DB: 12 Gaps: 1  
US-09-766-511B-53 (1-209) x US-10-090-466-3 (1-549)

```
QY 41 ValThrTyrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeu----- 57
|||:||||| |||:||||| ||||| ||||| |||||
Db 31 GTGCTTCAAAATTTTATGTATAGCAAACTGTCAAGAGGCTGTCCAAGTTACGAGAGTAT 90

QY 58 HisSerTyrHisSerLeuThrCysPheSerGluGlyThrLysValProAlaTrpGly 77
||||| ||||| ||||| ||||| ||||| |||||
Db 91 CAACAGTATCATYCAAGCCTGACCTCGTCATCGGAAGGAAGGACATAGAAGATTGGAGC 150

QY 78 CysCysProAlaSerTrpLysSerPheGlySerCysTyrPheIleSerSerGluGlu 97
||||| ||||| ||||| ||||| ||||| |||||
Db 151 TGCTGCCCCAACCCCTTGACATTCATTCAGTCTAGTCTACTTATTTCTACTGGGATG 210

QY 98 LysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValPhe 117
|||:||||| |||:||||| ||||| ||||| |||||
Db 211 CAATCTTGGACTAAGAGTCAAAAGAACTGTCTGTGATGGGGCTGATCTGTGTGTGATC 270

QY 118 AsnThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPhe 137
||||| ||||| ||||| ||||| ||||| |||||
Db 271 AACACCGAGGAAGAACAGGATTTTCATCTCAGAACTCGAAAGAAATTTCTTCTATTTT 330

QY 138 LeuGlyLeuSerAspProGlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrProTyr 157
||||| ||||| ||||| ||||| ||||| |||||
Db 331 CTGGGGCTGTCTCAGATCCAGGGGTCGGCGACATTTGGCAATGGTGGTTCACACACCATAC 390

QY 158 GluLysAsnValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAla 177
|||:||||| |||:||||| ||||| ||||| |||||
Db 391 AATGAAATGTACATCTTGGCAGTCAAGTGAACCCCAATAACCTTGATGAGCGTTGTGCG 450

QY 178 SerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArg 197
||| ||| ||| ||||| ||||| ||||| |||||
Db 451 ATAATAAATTTCCGTTCTTCAGAGAATGGGGCTGGAATGACATTCACCTGTCATGTACCT 510

QY 198 ArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209
:|: |||||:||||| ||||| |||||
Db 511 CAGAAGTCAATTTGCAAGATGAAGAAGATCTACATA 546
```

## RESULT 5

```
US-09-764-870-199
; Sequence 199, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ14
; CURRENT APPLICATION NUMBER: US/09/764.870
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 199
; LENGTH: 1091
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-870-199
```

Alignment Scores: 8.75e-43 Length: 1091  
Pred. No.: 458.00 Matches: 83  
Score: 62.57% Conservative: 34  
Percent Similarity: 44.39% Mismatches: 66  
Best Local Similarity: 39.31% Indels: 4  
Query Match: 10 Gaps: 3  
DB:

US-09-766-511B-53 (1-209) x US-09-764-870-199 (1-1091)

```
QY 26 IleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysValValThrTyrHisPhe 45
||| ||||| ||||| ||||| ||||| |||||
Db 172 ATATTTTCTCTGCTATTTGGCAATCTCATCTTTATTTGCTTTGTCATTTCTTTCAA--- 228
```

```
Qy 46 ThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHisSerTyrHisSerSerLeuThr 65
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 229 AAATATTCTCAGCTTCTTGAAAAAAGACTACAAAAAGAGCTGGTTTCATACAACTGGAG 288

Qy 56 CysPheSerGluGlyThrLysVal-----ProLalaTrpGlyCysCysProAlaSerTrp 83
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 289 TGTGTGAAAAAATAATGCCCGTGAAGAGACAGACCTGGAGCTGTTGCCAAAAAGAAATTGG 348

Qy 84 LysSerPheGlySerSerCysTyrPheIleSerSerGluGluLysValTrpSerLysSer 103
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 349 AAGTCATTAGTTCCTCACTGCTTATTCTACTCAATCAGCACTTGGCAAGACAGT 408

Qy 104 GluGlnAsnCysValGluMetGlyAlaHisLeuValPheAsnThrGluAlaGluGln 123
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 409 GAGAGGACTGTGTAGAATGGAGGCTCACCTGCTGTGATAAACACTCAAGAAGACGAG 468

Qy 124 AsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeuSerAspPro 143
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 469 GATTTCATCTCCAGAATCTGCAAGAAGAATCTGCTTATTATTGTTGGGGCTCTCAGATCCA 528

Qy 144 GlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsnValArgPhe 163
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 529 GAAGGTACGCACATTCGCAATGGTGTGATCAGACACCATACATGAATGAAGTTCACATTC 588

Qy 164 TrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAlaSerIleValPheTrpLys 183
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 589 TGGCATCCAGCTGAGCCAGTCATCCCAATCAGCGCTGCGTTGTCFAAATTTTCGTAATA 648

Qy 184 ---ProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArgAsnSerIleCys 202
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 649 TCACCCAAAAGATGGGCTGGAATGATGTTAATTGCTTGTCTCCTCAAGGTCAGTTTGT 708

Qy 203 GluMetAsnLysIleTyrLeu 209
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 709 GAGATGATGAAGATCCACTTA 729

RESULT 6
US-09-764-870-38
; Sequence 38, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT214
; CURRENT APPLICATION NUMBER: US/09/764,870
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
; LENGTH: 1096
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-870-38

Alignment Scores:
Pred. No.: 8,81e-43 Length: 1096
Score: 458.00 Matches: 83
Percent Similarity: 62.57% Conservative: 34
Best Local Similarity: 44.39% Mismatches: 66
Query Match: 39.31% Indels: 4
DB: 10 Gaps: 3

US-09-766-511b-53 (1-209) x US-09-764-870-38 (1-1096)

Qy 26 IleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysValValThrTyrHisPhe 45
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 182 ATATTTTCCGCTATTGGCAATCTCATCTTATTGCTTTTGCATTTCTTTCAA--- 238

Qy 46 ThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHisSerTyrHisSerSerLeuThr 65
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 239 AAATATTCTCAGCTTCTTGAAAAAAGACTACAAAAAGAGCTGGTGTTCATACAACTGGAG 298
```

```
Qy 66 CysPheSerGluGlyThrLysVal-----ProAlaTrpGlyCysCysProAlaSerTrp 83
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 299 TGTGTGAAAAAATAATGCCCGTGAAGAGACAGCCTGGAGCTGTGCCCAAAGAAATTGG 358

Qy 84 LysSerPheGlySerSerCysTyrPheIleSerSerGluGluLysValTrpSerLysSer 103
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 359 AAGTCATTAGTTCCTCACTGCTTATTCTACTGAATCAGCATCTTGGCAAGACAGT 418

Qy 104 GluGlnAsnCysValGluMetGlyAlaHisLeuValPheAsnThrGluAlaGluGln 123
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 419 GAGAGGACTGTGTAGAATGGAGGCTCACCTGCTGGTGATAAACACTCAAGAAGACGAG 478

Qy 124 AsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeuSerAspPro 143
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 479 GATTTCATCTCCAGAATCTGCAAGAAGAATCTGCTTATTATTGTTGGGGCTCTCAGATCCA 538

Qy 144 GlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsnValArgPhe 163
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 539 GAAGGTACGCAGCATTTGGCAATGGTGTGATCAGACACCATACAAATGAAAGTTCCACATTC 598

Qy 164 TrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAlaSerIleValPheTrpLys 183
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 599 TGGCATCCAGCTGAGCCAGTCATCCCAATGAGCGCTGCGTTGCTCTAAATTTTCGTAAA 658

Qy 184 ---ProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArgAsnSerIleCys 202
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 659 TCACCCAAAAGATGGGCTGGAATGATGTTAATTGCTTGTGCTCCTCAAGGTCAGTTTGT 718

Qy 203 GluMetAsnLysIleTyrLeu 209
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 719 GAGATGATGAAGATCCACTTA 739

RESULT 7
US-09-862-802-1
; Sequence 1, Application US/09862802
; Patent No. US20020165346A1
; GENERAL INFORMATION:
; APPLICANT: Schering-Plough Corporation
; TITLE OF INVENTION: MAMMALIAN MEMBRANE PROTEIN GENES; RELATED REAGENTS
; FILE REFERENCE: SF0695B
; CURRENT APPLICATION NUMBER: US/09/862,802
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/111,470
; PRIOR FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1104
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: mammalian nucleic acid
; NAME/KEY: CDS
; LOCATION: 242...952
; OTHER INFORMATION: protein coding sequence
US-09-862-802-1

Alignment Scores:
Pred. No.: 8,9e-43 Length: 1104
Score: 458.00 Matches: 83
Percent Similarity: 62.57% Conservative: 34
Best Local Similarity: 44.39% Mismatches: 66
Query Match: 39.31% Indels: 4
DB: 9 Gaps: 3

US-09-766-511b-53 (1-209) x US-09-862-802-1 (1-1104)

Qy 26 IleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysValValThrTyrHisPhe 45
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 395 ATATTTTCCGCTATTGGCAATCTCATCTTATTGCTTTTGTCTATTTCCTTCAA--- 451

Qy 46 ThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHisSerTyrHisSerSerLeuThr 65
```

```

Db 452 AAATATTCAGCTTCTTGGAAAAAAGACTACAAAGAGCTGGTTTCATACAACTGGAG 511
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 66 CysPheSerGluGlyThrLysVal-----ProAlaTrpGlyCysCysProAlaSerTrp 83
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 512 TGTGTGAAAAAATATGCGCTGGGAAGAGACAGCTGGAGCTGTGTCACCAAGAAATGG 571
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 84 LysSerPheGlySerSerCysTrpPheIleSerSerGluGlyValTrpSerLysSer 103
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 572 AAGTCAATTTAGTTCACACTGCTACTTATTTCTAGTCAATGACATCTTGGCAAGACAGT 631
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 104 GluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsnThrGluAlaGluGln 123
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 632 GAGAAGACTGCTAGAGTGGAGCTCACCTGCTGGTGATGAACACTCAAGAGAGCAG 691
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 124 AsnPheIleValGlnGlnLeuAsnGluSerPheSerTrpPheLeuGlyLeuSerAspPro 143
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 692 GATTTCATCTCCAGAATCTGCAAGAAGAATCTGCTTATTTGTGGGGCTCTCAGATCCA 751
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 144 GlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsnValArgPhe 163
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 752 GAAGGTGAGCGGACTTGGCAATGGGTGATCAGACACCATACAAATGAAGTTCCACATTC 811
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 164 TrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAlaSerIleValPheTrpLys 183
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 812 TGGCATCCAGCTGAGCCCACTGATCCCAATGAGCGCTGCTGCTAAATTTTCGTAA 871
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 184 ---ProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArgAsnSerIleCys 202
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 872 TCACCCAAAAGAGGGGCTGGAATGATGTTAATTGCTTGGCTCTCAAAAGTCAAGTTGT 931
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 203 GluMetAsnLysIleTyrLeu 209
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 932 GAGATGATGAAGATCCACTTA 952
      |||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
RESULT 8
US-09-833-381-119/c
; Sequence 119, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: ROBISON, Keith E.
; TITLE OF INVENTION: No. US20020132090A1e1 Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 119
; LENGTH: 1036
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(1036)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-119
Alignment Scores:
Pred. No.: 9,22e-43 Length: 1036
Score: 457.50 Matches: 89
Percent Similarity: 60.85% Conservative: 26
Best Local Similarity: 47.09% Mismatches: 58
Query Match: 39.27% Indels: 16
DB: 10 Gaps: 3
US-09-766-511b-53 (1-209) x US-09-833-381-119 (1-1036)
Qy 29 AlaLeuLeuSerAlaCysPheIleValSerCysValValThrTyrHisPheThrGly 48
      :|||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 701 AGCCTGAGGACCGAGTGCCTACAAATTTATGTATG----- 666
      :|||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :

```

```

Qy 49 GluThrGlyLysArgLeuSerGluLeuHisSerTyrHisSerSerLeu-----ThrCys 66
      :|||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 665 CAAACATCTCAGAGCTGCTCCAGTTACGGAGTTCAACGGTCTCACTCAAGCCGACCTGC 606
      :|||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 67 PheSerGluGlyThrLysValProAlaTrpGlyCysCysProAlaSerTrpLysSerPhe 86
      :|||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 605 GTCCTGTGAAGATAAGTACATAGAAGATTGGAGCTGCTGCCCAACCCCTTGGACTTCATTT 546
      :|||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 87 GlySerSerCysTrpPheIleSerSerGluLysValTrpSerLysSerGluGlnAsn 106
      :|||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 545 CAGCTAGTCTACTTATTTCTACCTGGGATGCAATCTTGGACTAAGACTCAAAAGAAC 486
      :|||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 107 CysValGluMetGlyAlaHisLeuValValPheAsnThrGluAlaGluGlnAsnPheIle 126
      :|||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 485 TGTCTGTGATGGGGCTGATCTGCTGATCAACACCAGGGAAGAACAGGATTTCATC 426
      :|||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 127 ValGlnGlnLeuAsnGluSerPheSerTrpPheLeuGlyLeuSerAspProGlnGlyAsn 146
      :|||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 425 ATTGAGATCTGAAAGAAATTTCTTATTTCTGGGGCTGTCAGATCCAGGGGGTCGG 366
      :|||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 147 AsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsnVal----- 161
      :|||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 365 CGACATGGCAATGGGTGACACACCATACATGAAATGTCAAGTATAGTAAT 306
      :|||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 162 -ArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAlaSerIleValPh 181
      :|||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 305 GAGATTCTGGCACCAGGTGAACCAATAACCTTGATGAGCGTTGTCGATAATAAATTT 246
      :|||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 181 eTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArgAsnSerIl 201
      :|||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 245 CCGTTCTTCAGAGAATGGGGCTGGAATGACATTCACATGTCATGTACCTCAGAAAGTCAAT 186
      :|||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 201 eCysGluMetAsnLysIleTyrLeu 209
      :|||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 185 TTGCAAGATGAAGAAGATCTACATA 161
      :|||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
RESULT 9
US-09-833-381-1339
; Sequence 1339, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: ROBISON, Keith E.
; TITLE OF INVENTION: No. US20020132090A1e1 Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1339
; LENGTH: 758
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-381-1339
Alignment Scores:
Pred. No.: 4,94e-41 Length: 758
Score: 440.50 Matches: 80
Percent Similarity: 61.38% Conservative: 36
Best Local Similarity: 42.33% Mismatches: 64
Query Match: 37.81% Indels: 9
DB: 10 Gaps: 5
US-09-766-511b-53 (1-209) x US-09-833-381-1339 (1-758)
Qy 23 ValAlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysValValThr 42
      :|||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 69 ATTGCTGATGTTTCTACTTCTTCTCAGTGTCTGTTTATTGCAAGTTGTTGGTACT 128
      :|||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Qy 43 TyrHis---PheThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHisSerTyrHis 61
      :|||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Db 129 CATCACAACTTTTTCACGCTGTAAAGAGGACACAGGAGTGCACAAGTTA---GAGCACCAT 185
      :|||   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :

```

QY	62	SerSerLeuThrCysPheSerGluGlyThrLysVal-----ProAlaTrp	76
		:::         :::::	
Db	186	GCAAGAGCTCAAAATGCATCAAAAGAAATCAGAACTTAAATCACTTAAGGAGACACCTGG	245
QY	77	GlyCysCysProAlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGlu	96
		:::::	
Db	246	AACTGTGTCTATTGACTGGAGAGCCTCCAGTCCAACTGCATTATTTCTCTTACTGAC	305
QY	97	GluLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValVal	116
		:::::	
Db	306	AACAAGACGTGGGTGAGAGTGAAAGAACTGTCAGGGATGGGGCCCATCTGATGACC	365
QY	117	PheAsnThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyr	136
		:::::     :::::	
Db	366	ATCAGACGGAAAGCTGACAGAACTTTATTTATTCAGTTCTGGATAGAGGGTTCCTAT	425
QY	137	PheLeuGlyLeuSerAspProGlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrPro	156
		:::::	
Db	426	TTCTCTGGACATTACAGATTAGAAATGCCAAAGGTCAGTGGCGTTGGGTGGACACAGCGCA	485
QY	157	TyrGluLysAsnValArgPheTrpHisLeuGlyGluProAsnHisSer---AlaGluGln	175
		:::     :::::	
Db	486	TTTAACCCACGACAGATTCTGGCATAGAATGAACCCGACAACCTCTCAGGGAGAAAAA	545
QY	176	CysAlaSerIleValPheTrpLysProThrGlyTrpAsnAspValIleCysGlu	195
		:::::	
Db	546	TGTGTGTCTTGTTTAT--AACCAAGATAAATGGGCCTGGNAATGATGTTCTCTTAAAC	602
QY	196	ThrArgArgAsnSerIleCysGluMet	204
		:::     :::::	
Db	603	TTTGAAAGCAAGTAGGATTTCTAAAAATA	629

```

RESULT 10
US-09-862-802-7
; Sequence 7, Application US/09862802
; Patent No. US20020165346A1
; GENERAL INFORMATION:
; APPLICANT: Schering-Plough Corporation
; TITLE OF INVENTION: MAMMALIAN MEMBRANE PROTEIN GENES; RELATED REAGENTS
; FILE REFERENCE: SFO695B
; CURRENT APPLICATION NUMBER: US/09/862,802
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/111,470
; PRIOR FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 1418
; TYPE: DNA
; ORGANISM: unknown
; FEATURE:
; OTHER INFORMATION: mammalian nucleic acid
; NAME/KEY: CDS
; LOCATION: 279...992
; OTHER INFORMATION: protein coding sequence
; NAME/KEY: misc_feature
; LOCATION: 1348
; OTHER INFORMATION: poly A addition motif
US-09-862-802-7

```

Alignment Scores:	
pred. No.:	4.1e-40
Score:	436.00
Percent Similarity:	61.03%
Best Local Similarity:	39.91%
Query Match:	37.42%
DB:	9
Length:	1418
Matches:	85
Conservative:	45
Mismatches:	73
Indels:	10
Gaps:	8

US-09-766-511B-53 (1-209) x US-09-862-802-7 (1-1418)

Oy 3 GlnGluGlnGlnProGlnSerThrGluLysArgGlyTrpLeuSerLeuArgLeuTrpSer 22

366	Db	AGAGAGAAACCTATCCGGTAGTCTAAGAAAGCGCTGGTTCCGCCCTCACTGCTTCTTACATCC	4253
23	Qy	ValAlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysValValThr	42
426	Db	CTGATG---CTACTTCTCTGCTGCTGGCAATCACATCTTAGTTGCTTTTATCATTT---	479
43	Qy	TyrHisPheThrTyrGlyGlu---ThrGlyLysArgLeuSerGluLeuHisSerTyrHis	61
480	Db	TATTTTCAAAAGTACTCTCAACTTCTTGAAGAAAAAAGCTGCAAAAAATATAATGCAC	539
62	Qy	SerSerLeuThrCysPheSerGluGlyThrLysVal-----ProAlaTrpGlyCysCys	79
540	Db	AATGAATGAAGTCAACAAAAAGTGTTCACCCATGGAGAACAAAGTCTGGAGCTGTTC	599
80	Qy	ProAlaSerTrpTyrSerPheGlySerSerCysTyrPheIleSer-----SerGluGlu	97
600	Db	CCAAAGGATTGGAGGCTATTGTGTTCCACTGCTACTGTGTTCCACAGTTCTTTCATCA	659
98	Qy	LysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValValPhe	117
660	Db	GCATCTTGGACAAGAGTGAAGAACTGCTCCCGCATGGGTGCTCATCTAGTGGTGATC	719
118	Qy	AsnThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPhe	137
720	Db	CAAGCCAGGAGACAGAGATTTCATCACTGGGATCTTGGACACTCATGCTGCTTATTTT	779
138	Qy	LeuGlyLeuSerAspProGlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrProTyr	157
780	Db	ATAGGGTTGTGGAT---ACAGGCCATCGGCAATGGCAATGGGTTCATCAGACACCATAT	836
158	Qy	GluLysAsnValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAla	177
837	Db	GAGAAAGTATCACATTCTGGCAACAATGGTGAAGCCAGCAGAGTGGCAATGAAAAATGTGC	896
178	Qy	SerIleValPhe---TrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThr	196
897	Db	ACAATAATTTACCGTTGGAAG---ACTGGATGGGCTGGAACGATATCTCTTCAGTCWT	953
197	Qy	ArgArgAsnSerIleCysGluMetAsnLysIleTyrLeu	209
954	Db	AAACAGAAGTCAGTTTGTCAAGTGAAGAAAAATAAATCTTA	992

```

RESULT 11
US-09-965-529-40
; Sequence 40, Application US/09965529
; Publication No. US20020182671A1
; GENERAL INFORMATION:
; APPLICANT: LAL, Preeti
; APPLICANT: YUE, Henry
; APPLICANT: TANG, Y. Tom
; APPLICANT: BANDMAN, Olga
; APPLICANT: BURFORD, Neil
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariyah R.
; APPLICANT: LO, Dyoung Aina M.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0731 USA
; CURRENT APPLICATION NUMBER: US/09/965,529
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/149,641; 60/164,203; PCT/US00/22315
; PRIOR FILING DATE:1999-08-17; 1999-11-09; 2000-08-14

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Inc
US-09-965-529-40

```

```
Alignment Scores:
Pred. No.:      4,23e-36      Length:      968
Score:          398.50      Matches:      84
Percent Similarity: 57.28%      Conservative: 34
Best Local Similarity: 40.78%      Mismatches: 81
Query Match:     34.21%      Indels:      7
DB:              9          Gaps:      5

US-09-766-511B-53 (1-209) x US-09-965-529-40 (1-968)
Qy 3 GlnGluGlnProGlnSerThrGluLysArgGlyTrpLeuSer-----LeuArgLeu 20
Db 174 AAATCATCTCAACACAAATGCACAGAG---AGAGGATGCTTCTCTCCCAAAATGTTCTTA 230
Qy 21 TrpSerValAlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysVal 40
Db 231 TGGACTGTTGCGGATCCCATCTTATTTCTCAGTGCCTGTTTCTATCACCAGATGTGT 290
Qy 41 ValThrTyrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHisSerTyr 60
Db 291 GTGACATTTCCGATC---TTTCAAACTGTGATGAGAAAAGTTTCAGCTACCTGAGAAT 347
Qy 61 HisSerSerLeuThrCysPheSerGluGlyThrLysValProAlaTrpGlyCysCysPro 80
Db 348 TTCACAGAGCTCTCTCTCTACAATATGATGCAGGT---TCAGTCAAGAAATTTCTGTCCA 404
Qy 81 AlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGluGluLysValTrp 100
Db 405 TTGAACCTGGGAATATTTTCAATCCAGCTGCTACTTCTTTTCTACTGACACCATTTCTCTGG 464
Qy 101 SerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsnThrGlu 120
Db 465 GCGTTAGTTAAAGAACTGCTCAGCCATGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 524
Qy 121 AlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeu 140
Db 525 GAGGAGCAGGAAATCTCTTCTCAAGAAACCTAAATGAGAGAGTTTATTTATTTGGACTG 584
Qy 141 SerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsn 160
Db 585 TCAGACCAGGTTGTCGAGGCTGAGTGGCAATGGGTGAGGCGGACACCTTTTGACAAAGTCT 644
Qy 161 ValArgPheTrpHisLeuGlyGluProAsnHisSerAla-----GluGlnCysAlaSer 178
Db 645 CTGAGCTCTCGGATGAGGGAGGCCCAACAACATAGCTACCTCGGAGGACTGTGCCACC 704
Qy 179 IleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArg 198
Db 705 ATGAGAGACTCTTCAAAACCAAGGCAAAATTTGAATGATGTAACTGTTTCTCCTCAATTAT 764
Qy 199 AsnSerIleCysGluMet 204
Db 765 TTTCGGATTGTGGAATG 782

RESULT 12
US-09-905-291A-376
; Sequence 376, Application US/09905291A
; Patent No. US20020160374A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
```

```
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905/291A
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 376
; LENGTH: 997
; TYPE: DNA
; ORGANISM: Homo Saplen
; US-09-905-291A-376

Alignment Scores:
Pred. No.:      4,42e-36      Length:      997
Score:          398.50      Matches:      84
Percent Similarity: 57.28%      Conservative: 34
Best Local Similarity: 40.78%      Mismatches: 81
Query Match:     34.21%      Indels:      7
DB:              9          Gaps:      5

US-09-766-511B-53 (1-209) x US-09-905-291A-376 (1-997)
Qy 3 GlnGluGlnProGlnSerThrGluLysArgGlyTrpLeuSer-----LeuArgLeu 20
Db 118 AAATCATCTCAACACAAATGCACAGAG---AGAGGATGCTTCTCTCCCAAAATGTTCTTA 174
Qy 21 TrpSerValAlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysVal 40
```





```
Db 409 GCGTTAGTTTAAAGAACTGCTCAGCCATGGGGCTCACCTGGTGGTTATCAACTCACAG 468
Qy 121 AlaGluGlnAsnPhelIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeu 140
Db 469 GAGGAGCAGGAATTCCTTCTCAAGAACAACCTAAATATGAGAGATTTTTTATTGGACTG 528
Qy 141 SerAspProGlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsn 160
Db 529 TCAGACCAGGTTCGCGAGGTTCAGTGGCAATGGGTGACGGCACACCTTTTGACAAAGTCT 588
Qy 161 ValArgPheTrpHisLeuGlyGluProAsnHisSerAla-----GluGlnCysAlaSer 178
Db 589 CTGAGCTTCTGGGATGAGGGAGGCCCAACAACATAGCTACCTCGGAGGACTGTGCCACC 648
Qy 179 IleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArg 198
Db 649 ATGAGAGACTCTCAACCCCAAGGCAAAATGGGAATGATGTAACCTGTTCCCTCAATTAT 708
Qy 199 AsnSerIleCysGluMet 204
Db 709 TTTCGGATTTCGGAATC 726

RESULT 14
US-09-766-511b-53 (1-209) x US-09-907-824-376 (1-997)
; Sequence 376, Application US/09907824
; Publication No. US20020197671A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Flivaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,824
; CURRENT FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: 09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
```

```
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 376
; LENGTH: 997
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-907-824-376

Alignment Scores:
Pred. No.: 4,42e-36 Length: 997
Score: 398.50 Matches: 84
Percent Similarity: 57.28% Conservative: 34
Best Local Similarity: 40.78% Mismatches: 81
Query Match: 34.21% Indels: 7
DB: Gaps: 5

US-09-766-511b-53 (1-209) x US-09-907-824-376 (1-997)
Qy 3 GlnGluGlnProGlnSerThrGluLysArgGlyTrpLeuSer-----LeuArgLeu 20
Db 118 AATCATCTCAACACACATGCACAGAG---AGAGATGCTTCTTCCCAAATGTTCTTA 174
Qy 21 TrpSerValAlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysVal 40
Db 175 TGGACTGTTGCTGGGATCCCATCTATTTCACAGTCCTGTTTCATCACCAGATGTGTT 234
Qy 41 ValThrTyrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHisSerTyr 60
Db 235 GTGACATTTCCGATC---TTTCAACCTGTGATGAGAAAAAGTTTCAGCTACCTGAGAAT 291
Qy 61 HisSerSerLeuThrCysPheSerGluGlyThrLysValProAlaTrpGlyCysCysPro 80
Db 292 TTCACAGAGCTCTCCCTGCTACAAATATGATCAGGT---TCAGTCAAGAAATTTGTGTCCA 348
Qy 81 AlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGluGluLysValTrp 100
Db 349 TTCAACTGGGAATATTTTCAATCCAGCTGCTACTTCTTTCTACTCACACCATTTCCCTGG 408
Qy 101 SerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsnThrGlu 120
Db 409 GCGTTAAGTTTAAAGAACTGCTCAGCCATGGGGCTCACCTGGTGGTTATCAACTCACAG 468
Qy 121 AlaGluGlnAsnPhelIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeu 140
Db 469 GAGGAGCAGGAATTCCTTCTCAAGAACAACCTAAATATGAGAGATTTTTTATTGGACTG 528
Qy 141 SerAspProGlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsn 160
Db 529 TCAGACCAGGTTCGCGAGGTTCAGTGGCAATGGGTGACGGCACACCTTTTGACAAAGTCT 588
Qy 161 ValArgPheTrpHisLeuGlyGluProAsnHisSerAla-----GluGlnCysAlaSer 178
Db 589 CTGAGCTTCTGGGATGAGGGAGGCCCAACAACATAGCTACCTCGGAGGACTGTGCCACC 648
```

QY 179 IleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArg 198  
Db 649 ATGAGAGACTCTTCAAAACCCAGGCAAAATTTGGAATGATGTAACCTGTTTCTCTCAATTAT 708  
QY 199 AsnSerIleCysGluMet 204  
Db 709 TTTTCGGATTGTGAAATG 726

RESULT 15

US-09-907-841-376  
; Sequence 376, Application US/09907841  
; Publication No. US20020198366A1  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc.  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Botstein, David  
; APPLICANT: Desnovers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, A.  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth, J.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Mather, Jennie P.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William, I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: 10466-14  
; CURRENT APPLICATION NUMBER: US/09/907,841  
; PRIOR APPLICATION NUMBER: PCT/US00/04414  
; PRIOR FILING DATE: 2000-02-22  
; PRIOR APPLICATION NUMBER: US 60/143,048  
; PRIOR FILING DATE: 1999-07-07  
; PRIOR APPLICATION NUMBER: US 60/145,698  
; PRIOR FILING DATE: 1999-07-26  
; PRIOR APPLICATION NUMBER: US 60/146,222  
; PRIOR FILING DATE: 1999-07-28  
; PRIOR APPLICATION NUMBER: PCT/US99/20594  
; PRIOR FILING DATE: 1999-09-08  
; PRIOR APPLICATION NUMBER: PCT/US99/20944  
; PRIOR FILING DATE: 1999-09-13  
; PRIOR APPLICATION NUMBER: PCT/US99/21090  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR APPLICATION NUMBER: PCT/US99/21547  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR APPLICATION NUMBER: PCT/US99/23089  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR APPLICATION NUMBER: PCT/US99/28214  
; PRIOR FILING DATE: 1999-11-29  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 423  
; SEQ ID NO 376  
; LENGTH: 997  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-09-907-841-376

Alignment Scores: 4.42e-36 Length: 997  
Pred. No.:

Score: 398.50 Matches: 84  
Percent Similarity: 57.28% Conservatives: 34  
Best Local Similarity: 40.78% Mismatches: 81  
Query Match: 34.21% Indels: 7  
DB: 9 Gaps: 5  
US-09-766-511b-53 (1-209) x US-09-907-841-376 (1-997)  
QY 3 GlnGluGlnProGlnSerThrGluLysArgGlyTrpLeuSer-----LeuArgLeu 20  
Db 118 AAATCATCTGAAACACAATGCACAGAG--AGAGGATGCTTCTTCTCCCAAAATGTTCTTA 174  
QY 21 TrpSerValAlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysVal 40  
Db 175 TGGACTGTTGCTGGATCCCATCTTATTTCTCAGTGCCTGTTTTCATCACCAGATGTT 234  
QY 41 ValThrTyrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHisSerTyr 60  
Db 235 GTGACATTTCCGATC---TTTCAAACCTGTGATGAGAAAAGTTTCAGCTACTGAGAAT 291  
QY 61 HisSerSerLeuThrCysPheSerGluGlyThrLysValProAlaTrpGlyCysCysPro 80  
Db 292 TTCACAGAGCTCTCCIGCTACAATTATGATCAGGT---TCAGTCAAGAATTTGTTGCCA 348  
QY 81 AlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSerGluGluLysValTrp 100  
Db 349 TTGAACCTGGGAATATTTTCAATCCAGCTGCTACTTTCTTTCTACTGACACCATTTCTCG 408  
QY 101 SerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValPheAsnThrGlu 120  
Db 409 GCCTTAAGTTTAAAGAACTGCTCAGCCATGGGGGCTCACCTGGTGTATATCAACTCACAG 468  
QY 121 AlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeu 140  
Db 469 GAGGAGCAGGAATTCCTTCCACAAAGAACTTAAATGAGAGAGATTTTATTGGACTG 528  
QY 141 SerAspProGlnGlyAsnAsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsn 160  
Db 529 TCAGACCAGGTTGTCGAGGGTCAGTGGCAATGGGTGGAGGCACACCTTTTGACAAAGTCT 588  
QY 161 ValArgPheTrpHisLeuGlyGluProAsnHisSerAla-----GluGlnCysAlaSer 178  
Db 589 CTGAGCTTCTGGATGTAGGGAGGCCCAACACATAGCTACCCCTGGAGGACTGTGCCACC 648  
QY 179 IleValPheTrpLysProThrGlyTrpGlyTyrAsnAspValIleCysGluThrArgArg 198  
Db 649 ATGAGAGACTCTTCAAAACCCAGGCAAAATTTGGAATGATGTAACCTGTTTCTCTCAATTAT 708  
QY 199 AsnSerIleCysGluMet 204  
Db 709 TTTTCGGATTGTGAAATG 726

Search completed: February 20, 2003, 20:41:58  
Job time : 62 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 20, 2003, 19:24:22 : Search time 1689 Seconds  
(without alignments)  
2004.060 Million cell updates/sec

Title: US-09-766-511b-53  
Perfect score: 1165  
Sequence: 1 MMQEQQPQSTKRGWLSRL.....NDVICETRRNSICEMNKIYL 209

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-Q=/cgn2\_1/USPTO.spool/US09766511/runat\_19022003\_094339\_29300/app\_query.fasta\_1.391  
-DB=EST -QFMT=fastap -SUFFIX=1st -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosom62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09766511@cgn\_1.1.899\_@runat\_19022003\_094339\_29300 -NCPU=3  
-NO\_XLPXY -NO\_MMALP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMECUT=120  
-WARN\_TIMECUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST.\*  
1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estin.\*  
4: em\_estmu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_estt.\*  
9: gb\_est1.\*  
10: gb\_est2.\*  
11: gb\_est3.\*  
12: gb\_est4.\*  
13: gb\_est5.\*  
14: gb\_est6.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: gb\_gss.\*  
18: em\_gss\_hum.\*  
19: em\_gss\_inv.\*  
20: em\_gss\_pin.\*  
21: em\_gss\_vrt.\*  
22: em\_gss\_fun.\*  
23: em\_gss\_mam.\*  
24: em\_gss\_mus.\*  
25: em\_gss\_other.\*  
26: em\_gss\_pro.\*  
27: em\_gss\_rod.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	694.5	59.6	971	13	BI696642	BI696642 603348532
2	682	58.5	504	12	BF603486	BF603486 269014 MA
3	666	57.2	470	12	BF774158	BF774158 283735 MA
4	475.5	40.8	631	11	AK007794	AK007794 Mus muscu
c 5	439.5	37.7	582	9	AI183967	AI183967 qd69e10.x
c 6	427.5	36.7	568	10	BE200183	BE200183 ug60b08.x
7	419	36.0	514	12	BF152383	BF152383 u232f08.y
8	419	36.0	558	9	AI391056	AI391056 mcl3h05.y
9	414.5	35.6	812	11	AK020363	AK020363 Mus muscu
c 10	400.5	34.4	403	10	AW822082	AW822082 up67g10.x
c 11	390.5	33.5	1200	11	AK014135	AK014135 Mus muscu
c 12	382	32.8	1013	9	AA868502	AA868502 ak43b01.s
c 13	377.5	32.4	753	9	AI949910	AI949910 wq04a07.x
14	375.5	32.2	561	10	BE650539	BE650539 UT-M-BH2
15	373.5	32.1	1032	10	BE311916	BE311916 601143328
16	363	31.2	715	10	AV716802	AV716802 AV716802
17	362	31.1	889	12	BG164873	BG164873 602343809
18	359	30.8	676	13	BI909298	BI909298 603067260
c 19	352	30.2	831	12	BG065735	BG065735 H3034F05-
20	350.5	30.1	582	10	AV716725	AV716725 AV716725
21	349.5	30.0	631	12	BG080418	BG080418 H3053D08-
22	348.5	29.9	498	10	BE689245	BE689245 uw51d04.y
23	342.5	29.4	660	10	BE292395	BE292395 601057892
24	342	29.4	1007	14	BQ715169	BQ715169 AGENCOURT
25	333.5	28.6	854	13	BI910606	BI910606 603069821
c 26	316.5	27.2	667	10	BE688920	BE688920 uw51d04.x
27	315	27.0	576	12	BG079007	BG079007 H3034F05-
28	314.5	27.0	334	10	BE690187	BE690187 uw65c04.y
c 29	314.5	27.0	738	14	BQ004812	BQ004812 UI-H-E10-
c 30	299.5	25.7	448	9	AA46401	AA46401 zw58d11.r
c 31	290.5	24.9	666	10	AW242656	AW242656 xm89a06.x
c 32	288.5	24.8	681	12	BF020383	BF020383 uw65c04.x
c 33	278.5	23.9	678	12	BF227688	BF227688 u232f08.x
c 34	267.5	23.0	559	10	BE120531	BE120531 UI-R-CA0-
35	267	22.9	432	10	BB863598	BB863598 BB863598
36	261	22.4	604	10	AW743449	AW743449 up67g10.y
37	259	22.2	645	13	BI328284	BI328284 602985590
38	259	22.2	714	12	BF533278	BF533278 602073733
c 39	254.5	21.8	366	13	BI027634	BI027634 CW3-WT029
40	251.5	21.6	458	10	BE632986	BE632986 uv71e02.y
c 41	251.5	21.6	816	14	BQ018186	BQ018186 UI-H-DT1-
42	249	21.4	668	14	BQ674807	BQ674807 AGENCOURT
43	249	21.4	785	14	BQ648893	BQ648893 AGENCOURT
44	249	21.4	886	14	BQ651695	BQ651695 AGENCOURT
c 45	249	21.4	914	9	AL564408	AL564408 AL564408

ALIGNMENTS

RESULT 1  
BI696642  
LOCUS 603348532F1 NC1\_CGAP\_Mam2 Mus musculus cDNA clone IMAGE:5376247 5',  
DEFINITION BI696642 mRNA sequence.  
ACCESSION BI696642  
VERSION BI696642.1 GI:15659271  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 971)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)



```
Db 7 GCTGAAGTGCACACACACACACATTAAATCTAACCTGCTTACAGTGAAGGACAAAGGTCACA 66
Qy 75 Ala-----TTPGlyCysCysProAlaSerTrpLysSerPheGlySerSerCysTrpH 92
Db 67 NGAAGAAGATTGGGGATGTTGCCACAGTACCTGCAAGCCGTTTGGTCCAGCTGCTACTT 126
Qy 92 eIleSerSerGluGluLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAl 112
Db 127 TATTCTTCTGAAGAGAATTCTGGGCTAAGAGTGAAGAGTGCACAGAACTGGCATGGGAGC 186
Qy 112 aHisLeuValValPheAsnThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGl 132
Db 187 TCACCTGGTGTGATCAACACAGAAACAGACAGAGGATTCATTATCCAGCAGCTGAATAA 246
Qy 132 uSerPheSerTrpPheLeuGlyLeuSerAspProGlnGlnAsnAsnTrpGlnTrpIl 152
Db 247 AACATTTCTTATTCTTGGGACTCTCAGACCCCAAGGAATGGCAACTGGCAATGGAT 306
Qy 152 eAspLysThrProTyrGluLysAsnValArgPheTrpHisLeuGlyGluProAsnHisSe 172
Db 307 TGATCAGACACCTTACAAGGAAATGTGAGATTTTGGCCCAAAATGAACCCAACTTTTC 366
Qy 172 rAlaGluGlnCysAlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspVa 192
Db 367 TCACAGGAATGTGCTTCAGTTGTTTCTGGGATGGGAGAGGATGGGGCTGGAATGATG 426
Qy 192 lIleCysGluThrArgArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209
Db 427 TTTCTGTGATTTCTAAAGGAATCAATATGTGAGATGAAGAAGATTTACCTA 478

RESULT 3
BF774158 470 bp mRNA linear EST 25-APR-2001
LOCUS 283735 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION BF774158
ACCESSION BF774158
VERSION BF774158.1 GI:12122058
KEYWORDS EST.
SOURCE cow.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 470)
AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
,G.L., Heaton,M.P., Laegreid,W., Rohrer,G.A., Chitko-McKown,C.G.,
Perlea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
Keele,J.W.
TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
21180013
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR primers
FORWARD: AGGAACACGCTATGACCAT
BACKWARD: GTTTCACGTCAGTCAGC
Plate: 80 row: M column: 16
Seq primer: ATTTAGTGACACTATAG.
Location/Qualifiers
1. 470
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 3BOV"
/tissue_type="pooled"
/lab_host="DH10B"
```

```
/note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
Library made from pooled tissue from marrow, alveolar
macrophage, ovary, fetal semitendinosus muscle, and fetal
longissimus muscle."
BASE COUNT 138 a 84 c 122 g 124 t 2 others
ORIGIN
Alignment Scores:
Pred. No.: 2e-66 Length: 470
Score: 666.00 Matches: 116
Percent Similarity: 85.99% Conservative: 19
Best Local Similarity: 73.89% Mismatches: 18
Query Match: 57.17% Indels: 4
DB: 12 Gaps: 1
US-09-766-511B-53 (1-209) x BF774158 (1-470)
Qy 52 LysArgLeuSer-GluLeuHisSerTyrHisSerSerLeuThrCysPheSerGluGlyTh 71
Db 1 AAGAGGCTGTCNTGAACCTGCACACACACCACTTCAATCAACCTGCTTCAGTGAAGGAC 60
Qy 71 rLysValProAla-----TTPGlyCysCysProAlaSerTrpLysSerPheGlySerS 89
Db 61 AAGGGTGACANGAAGATTGGGGATGTTGCCAGTACTCTGGAAGCCGCTTGGTTCCA 120
Qy 89 eCysTyrPheIleSerSerGluLysValTrpSerLysSerGluGlnAsnCysVal 109
Db 121 GCTGCTACTTATTCTCTGAAGAGAAATTTCTGGCTAAGAGTGAAGCAACTGCATTG 180
Qy 109 LuMetGlyAlaHisLeuValPheAsnThrGluAlaGluGlnAsnPheIleValGlnG 129
Db 181 GGATGGAGCTCACTTGGTGTATCAACACAGAACAGAGAGATTTTCATTATCCAGC 240
Qy 129 InLeuAsnGluSerPheSerTyrPheLeuGlyLeuSerAspProGlnGlnAsnAsn 149
Db 241 AGCTGAATAAAACATTTTCTTATTTTCTGGACTCTCAGACCCCAAGGAATGGCAACT 300
Qy 149 rpGlnTrpIleAspLysThrProTyrGluLysAsnValArgPheTrpHisLeuGlyGlu 169
Db 301 GGCAATGGATTGATCAGACACCTTTACAAGGAAATGTGCAGATTTTGGCACCACAA 360
Qy 169 roAsnHisSerAlaGluGlnCysAlaSerIleValPheTrpLysProThrGlyTrpGly 189
Db 361 CCAACTTTTTCGACAGAAATGTCTTCAGTTGTTTCTGGGATGGGAGAGGAGGGGCT 420
Qy 189 rpAsnAspValIleCysGluThrArgArgAsnSerIleCysGluMet 204
Db 421 GGAATGATGTTTCTGTGATTTCTAAAGGAATCAATATGTGAGATG 467

RESULT 4
AK007794 631 bp mRNA linear HTC 19-JAN-2002
LOCUS AK007794
DEFINITION Mus musculus 10 day old male pancreas cDNA, RIKEN full-length
enriched library, clone:1810046124; similar to DENDRITIC CELL
IMMUNORECEPTOR, full insert sequence.
ACCESSION AK007794
VERSION AK007794.1 GI:12841567
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (Strain:C57BL/6J) 10 day old male pancreas cDNA to
mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library
clone:1810046124.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
```

## TITLE

JOURNAL Prepare full-length cDNA libraries for rapid discovery of new genes  
MEDLINE Genome Res. 10 (10), 1617-1630 (2000)  
PUBMED 20499374  
REFERENCE 11042159

## AUTHORS

3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,  
Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,  
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,  
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,  
Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,  
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,  
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multipillar sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861

## TITLE

JOURNAL Prepare full-length cDNA libraries for rapid discovery of new genes  
MEDLINE Genome Res. 10 (11), 1757-1771 (2000)  
PUBMED 20530913  
REFERENCE 11076861

## AUTHORS

4 Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,  
Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,  
Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,  
Saito,T., Okazaki,Y., Gojohori,T., Bono,H., Kasukawa,T., Saito,R.,  
Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,  
Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,  
Kuehl,P., Lewis,S., Matsuo,Y., Nikaide,I., Pesole,G.,  
Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,  
Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,  
Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,  
Carninci,P., De Bonaldo,M.F., Brownstein,M.J., Bult,C.,  
Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,  
Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,  
Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,  
Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,  
Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,  
Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,  
Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohsaki,S.,  
and Hayashizaki,Y.

JOURNAL Functional annotation of a full-length mouse cDNA collection  
MEDLINE Nature 409 (6821), 685-690 (2001)  
PUBMED 21085660  
REFERENCE 11217851

## AUTHORS

5 (bases 1 to 631)  
Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,  
Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C.,  
Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T.,  
Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F.,  
Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T.,  
Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S.,  
Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K.,  
Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J.,  
Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,  
Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,  
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,  
Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I.,  
Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and  
Hayashizaki,Y.

## Direct Submission

Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,  
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,  
Fax:81-45-503-9216)

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for  
further details.

cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues. First strand cDNA was primed with a primer  
[5' GAGAGAGAGAGGATCCAGAGCTCTTTTTTTTTTTVN 3'], cDNA was

prepared by using trehalose thermo-activated reverse transcriptase  
and subsequently enriched for full-length by cap-trapper. cDNA went  
through one round of subtraction to Rot = 20.0. Second strand cDNA  
was prepared with the primer adapter of sequence[5'  
GAGAGAGAGCGCGCAATTAACTTCGAGTTAAATAATCCCCCCCCC 3'], cDNA  
was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3',  
end: SstI. Host: SOLR.

## FEATURES

Location/Qualifiers  
Source

1..631  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="FANTOM\_DB:1810046124"  
/db\_xref="MGI:1901987"  
/db\_xref="taxon:10090"  
/clone="1810046124"  
/sex="male"  
/tissue\_type="pancreas"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="10 day old"  
94..624

## CDS

/note="data source:SPTR, source key:Q9QZL5, evidence:ISS  
putative  
similar to DENDRITIC CELL IMMUNORECEPTOR"  
/codon\_start=1  
/protein\_id="BAB25260.1"  
/db\_xref="GI:12841568"  
/db\_xref="MGI:1917060"  
/translation="MVGRQLQKAVSWLSRLMSAAVSIILLSTCTFASCVKVMSC  
CPDKMLFGSHCYLPTVFSSASNKSENCRRMGAHLVHHSQEQDITGLDIHA  
AYFTGLMDTGHQWQWDDTPYEESTFWHNGEPSSDNEKCVTVYRRNIGWWDIS  
CNLKQKSVCCOMKINL"

BASE COUNT 174 a 129 c 157 g 171 t  
ORIGIN

Alignment Scores:  
Pred. No.: 2,76e-44 Length: 631  
Score: 475.50 Matches: 92  
Percent Similarity: 57.82% Conservative: 30  
Best Local Similarity: 43.60% Mismatches: 52  
Query Match: 40.82% Indels: 37  
DB: 11 Gaps: 6

US-09-766-511b-53 (1-209) x AK007794 (1-631)

Qy 1 MetMetGlnGlnGlnProGlnSerThrGlyLysArgGlyTrpLeuSerLeuArgLeu 20

Db 94 ATGCTTCAGGAAGACAGCTACAAAGG---AAGCAGCTCTCTGG---TCCCTGAGACTC 147

Qy 21 TrpSerValAlaGlyIleSerIleAlaLeuLeuSerAlaCysPheIleValSerCysVal 40

Db 148 TGGTCAGCTGCTGTGATTTCTATCTACTCTCAGCACCTGTTTCATTGCGAGCTGTGTG 207

Qy 41 ValThrTyrHisPheThrTyrGlyGluThrGlyLysArgLeuSerGluLeuHisSerTyr 60

Db 207 ----- 207

Qy 61 HisSerSerLeuThrCysPheSerGluGlyThrLysValProAlaTrpGlyCysCysPro 80

Db 208 -----GACAAAGTC-----TGGAGCTGTGCCCCA 231

Qy 81 AlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSer-----GluGluLys 98

Db 232 AAGGATTGGGAAGTATTGGTTCCTACTGTTCCCATGCTACTGGTTCCTCCACAGCA 291

Qy 99 ValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuValValPheAsn 118

Db 292 TCITGGACAAAGAGTGGAGAGAACTGCTCCCATGGGTGCTCATCTAGTGTGTATCCAT 351

Qy 119 ThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSerTyrPheLeu 138

Db 352 AGCCAGGAGGAGGAGGAGTTCATCATCTGGGATCTTGGACATTCATGCTCTTATTTATA 411

Qy 139 GlyLeuSerAspProGlnGlyAsnAsnAsnThrGlnTrpIleAspLysThrProTyrGlu 158

```

Db 412 GGGTGTGGGAT---ACAGGCCATCGCAATGGCAATGGTTGATCAGACACCATATGAA 468
QY 159 LysAsnValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGlnCysAlaSer 178
Db 469 GAAAGTGTCACTCTGGCACAATGGTGAGCCCGCAGCAGTGACAAATGAAAAATGTGTACA 528
QY 179 IleValPheTrpLysProThrGlyTrpGlyTrpAsnAspValIleCysGluThrArgArg 198
Db 529 GTATATTACCGTCGGAATATTGGATGGGCGTGAATGATATCTCTTGCAATCTTAAACAG 588
QY 199 AsnSerIleCysGluMetAsnLysIleTyrLeu 209
Db 589 AAGTCAGTTTGTGAGATGAAGAAAAATAACTTA 621

RESULT 5
AL183967/c
LOCUS      AL183967               582 bp      mRNA      linear      EST 09-OCT-1998
DEFINITION qd69e10.x1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1734762
              3' similar to SW:LECH_MOUSE P34927 ASIALOGLYCOPROTEIN RECEPTOR 1 ;,
              mRNA sequence.
ACCESSION   AL183967
VERSION     AL183967.1 GI:3734605
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 582)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
              Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-r@mail.nih.gov
              cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
              , Ph.D.
              cDNA Library Arrayed by: Greg Lennon, Ph.D.
              DNA Sequencing by: Washington University Genome Sequencing Center
              Clone distribution: NCI-CGAP clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              www-bio.llnl.gov/bbrp/image/image.html
              Seq primer: -400P from Glibco
              High quality sequence stop: 453.
FEATURES    Location/Qualifiers
             source
               1..582
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone="IMAGE:1734762"
               /clone_lib="Soares_testis_NHT"
               /sex="male"
               /lab_host="DH10B"
               /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
               polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
               was prepared from mRNA obtained from Clontech Laboratories
               , Inc., and primed with a Not I - oligo(dT) primer [5',
               TGTTACCAATCTGAGTGGGAGCGGCCCAATTTTTTTTTTTT 3'].
               Double-stranded cDNA was ligated to Eco RI adaptors
               (Pharmacia), digested with Not I and cloned into the Not I
               and Eco RI sites of the modified pT7T3 vector. Library
               went through one round of normalization to Cot5, and was
               constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 175 a 131 c 105 g 171 t
ORIGIN
```

```

US-09-766-511b-53 (1-209) x AI183967 (1-582)
QY 76 TrpGlyCysCysProAlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSer 95
Db 579 TGGACTGCTGCCCAACCCCTGGACTTCATTTCAGTCTAGTGTCTATTATTCTACT 520
QY 96 GluGluLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuVal 115
Db 519 GGGATGCAATCTGGCAGTAAGAGTCAAAAGAACTGTTCTGTGATGGGGCTGATCTGGTG 460
QY 116 ValPheAsnThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGluSerPheSer 135
Db 459 GTGATCAACACACGAGGAAGACAGGATTTCATTCATTCAGAAATCTGAAAAGAAATTTCTTCT 400
QY 136 TyrPheLeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThr 155
Db 399 TATTTTCTGGGCGTGTGAGATCCAGGGGTGGCCACATTTGCAATGGTGTGACCGACA 340
QY 156 ProTyrGluLysAsnVal-----ArgPheTrpHisLeuGlyGluProAs 170
Db 339 CCATCAATGAAAAATGTCACGTGAGTATAGAATGAGATTCTGGCAGTCAGGTGAACCCAA 280
QY 170 nHisSerAlaGluGlnCysAlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAs 190
Db 279 TAACCTTGATGAGCGTTGTGCGAATAATAATTTCCGTTCTTCAGAGAAGATGGGGCTGGA 220
QY 190 nAspValIleCysGluThrArgArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209
Db 219 TGACATTCTACTGTCATGTACCTCAGAAAGTCAATTTGCAAGATGAAGAAGATCTACATA 162

RESULT 6
BE200183/c
LOCUS      BE200183               568 bp      mRNA      linear      EST 26-JUN-2000
DEFINITION ug60b08.x1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
              IMAGE:1546743 3' similar to TR:Q9Z2H6 09Z2H6 C-TYPE LECTIN. ;, mRNA
              sequence.
ACCESSION   BE200183
VERSION     BE200183
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 568)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
              Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Other_ESTs: ug60b08.y1
              Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-r@mail.nih.gov
              This clone is available royalty-free through LLNL ; contact the
              IMAGE Consortium (info@image.llnl.gov) for further information.
              MGI:952091
              High quality location sequence stop: 464.
FEATURES    Location/Qualifiers
             source
               1..568
               /organism="Mus musculus"
               /db_xref="taxon:10090"
               /clone="IMAGE:1546743"
               /clone_lib="Soares_mammary_gland_NMLMG"
               /sex="female (lactating)"
               /tissue_type="mammary gland"
               /lab_host="DH10B"
               /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
               polylinker; 1st strand cDNA was prepared from mammary
               gland tissue from a lactating female, and was then primed
               with a Not I - oligo(dT) primer. Double-stranded cDNA was
               ligated to Eco RI adaptors (Pharmacia), digested with Not
               I and cloned into the Not I and Eco RI sites of the
               modified pT7T3 vector. Library is normalized. Library
               was constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 155 a 128 c 110 g 175 t
```





**DEFINITION**  
mcl3h05.y1 Soares mouse p3NMF19.5 Mus musculus cDNA clone  
IMAGE:348441 5' similar to SW:MANR\_HUMAN P22897 MACROPHAGE MANNOS  
RECEPTOR PRECURSOR. ; mRNA sequence.

**ACCESSION**  
AI391056

**VERSION**  
AI391056.1 GI:4217063

**KEYWORDS**  
EST.

**SOURCE**  
house mouse

**ORGANISM**  
Mus musculus

**REFERENCE**  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 558)

**AUTHORS**  
Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,  
Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person  
, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter  
, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,  
Waterston, R., and Wilson, R.

**TITLE**  
The WashU-NCI Mouse EST Project 1999

**JOURNAL**  
Unpublished (1999)

**COMMENT**  
Contact: Marra M/WashU-NCI Mouse EST Project 1999  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@wustl.edu  
This clone is available royalty-free through LNL ; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
This read is a RESEQUENCE of a previously sequenced mouse clone  
This read has been verified (found to hit its original self in the  
correct orientation)  
MGI:220241  
Seq primer: -40RP from Gibco  
High quality sequence stop: 491  
POLYA=No.

**FEATURES**

**Source**  
1. .558  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="IMAGE:348441"  
/clone\_lib="Soares mouse p3NMF19.5"  
/dev\_stage="19.5 dpc total fetus"  
/lab\_host="DH10B (ampicillin resistant)"  
/note="Vector: pT73D (Pharmacia) with a modified  
polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
was primed with a Not I - oligo(dT) primer [5',  
TCCTACCAATCGAAGTGGAGCGCGCATTTTGTGTTTGT 3'],  
double-stranded cDNA was size selected, ligated to Eco RI  
adapters (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of a modified pT73 vector  
(Pharmacia). Library went through one round of  
normalization to a Cot = 5. Library constructed by Bento  
Soares and M.Fatima Bonaldo. RNA was kindly provided by  
Dr. Minoru Ko (Wayne State University)."

**BASE COUNT**  
171 a 108 c 129 g 150 t

**ORIGIN**

**Alignment Scores:**

Pred. No.:	7,33e-38	Length:	558
Score:	419.00	Matches:	73
Percent Similarity:	64.97%	Conservative:	29
Best local Similarity:	46.50%	Mismatches:	53
Query Match:	35.97%	Indels:	2
DB:	9	Gaps:	2

US-09-766-511B-53 (1-209) x AI391056 (1-558)

**Qy** 54 LeuSerGluLeuHisSerTyrHisSerSerLeuThrCysPheSerGluGlyThrLysVal 73  
::: |||||:::  
||| ::::

**Db** 14 ATAAAGAACTGAACATACTAATGGAGTGATACAAAATGGGCTTCACTCTTGAAGAC 73

**Qy** 74 ProAlaTrpGlyCysCysProAlaSerTrpLysSerPheGlySerSerCysTyrPheIle 93  
||| |||||:::  
||| |||||:::

**Db** 74 AAAGTCGGAGCTGTTGCCCAAGGATGGAGCCGCTTTGGTTCTTCACTGCTACTTCACT 133

**Qy** 94 SerSerGlu---GluLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAla 112  
|||:::|:::  
|||:::|:::

**Db** 134 TCAACTGACTTGGTGGCATCTTGCATGAGATGAAGGAACACTGCTTCCACATGGTGCT 193

**Qy** 113 HisLeuValValPheAsnThrGluAlaGluGlnAsnPheIleValGlnGlnLeuAsnGlu 132  
||||||:::  
||||||:::

**Db** 194 CATCTGGTGGTGCATCCAGCAGCAGGAACAGGATTTTCATCACTGGGACTCTGGACACT 253

**Qy** 133 SerPheSerTyrPheLeuGlyLeuSerAspProGlnGlyAsnAsnAsnTrpGlnTrpIle 152  
|||:::|:::  
|||:::|:::

**Db** 254 GGTACTGCTTATTTATAGGACTTTTCAATCCA---GGTGATCAACAATGGCAATGGATT 310

**Qy** 153 AspLysThrProTyrGluLysAsnValArgPheTrpHisLeuGlyGluProAsnHisSer 172  
|||:::|:::  
|||:::|:::

**Db** 311 GATCAGACACCGTACGATGATAATACCACATCTCTGGCACAAGGTGAGCCCTAGCAGTGAC 370

**Qy** 173 AlaGluGlnCysAlaSerIleValPheTrpLysProThrGlyTrpGlyTrpAsnAspVal 192  
||||||:::  
|||:::|:::

**Db** 371 AATGAACAGTGTGTATATAATAATCATGTCAGAGTACTGGATGGGCTGGAGTGATATC 430

**Qy** 193 IleCysGluThrArgArgAsnSerIleCysGluMetAsnLysIleTyrLeu 209  
|||:::|:::  
|||:::|:::

**Db** 431 CTTGCACTGATAACAGAACTCAATTGCCATGTGAAAAAATATACTTA 481

**RESULT** 9  
AK020363

**LOCUS**  
AK020363

**DEFINITION**  
Mus musculus adult male epididymis cDNA, RIKEN full-length enriched library, clone:9230119016:C-type (calcium dependent, carbohydrate recognition domain) lectin, superfamily member 6, full insert sequence.

**ACCESSION**  
AK020363

**VERSION**  
AK020363.1 GI:12860929

**KEYWORDS**  
HTC; CAP trapper.

**SOURCE**  
Mus musculus (strain:C57BL/6J) adult male epididymis cDNA to mRNA, clone:lib:RIKEN full-length enriched mouse cDNA library  
clone:9230119016.

**ORGANISM**  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

**REFERENCE**  
1 Carninci,P. and Hayashizaki,Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
99279253  
10349636

**REFERENCE**  
2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
11042159

**REFERENCE**  
3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,K., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipillar sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861

**REFERENCE**  
4 Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I., Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R., Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,





**TITLE** Direct Submission  
**JOURNAL** Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Gemomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)  
**COMMENT** Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.  
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Gemomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAAGATCCAGAGCTCTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 50.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATCTCGAGTTAATAATTAATCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.

**FEATURES**  
 source  
 1..1200  
   /organism="Mus musculus"  
   /strain="C57BL/6J"  
   /db\_xref="FANTOM\_DB:3110037K17"  
   /db\_xref="MGD:MGI:1906103"  
   /db\_xref="taxon:10090"  
   /clone="3110037K17"  
   /tissue\_type="head"  
   /clone\_lib="RIKEN full-length enriched mouse cDNA library"  
   /dev\_stage="13 days embryo"  
 misc\_feature  
 1..1200  
   /note="data source:SPTR, source key:O9Q2I5, evidence:ISS similar to DENDRITIC CELL IMMUNORECEPTOR"  
   /db\_xref="MGD:MGI:1920399"

BASE COUNT	349 a	251 c	235 g	365 t
ORIGIN				

Alignment Scores:  
 Pred. No.: 4.65e-34 Length: 1200  
 Score: 390.50 Matches: 77  
 Percent Similarity: 58.70% Conservative: 31  
 Best Local Similarity: 41.85% Mismatches: 71  
 Query Match: 33.52% Indels: 6  
 DB: 11 Gaps: 3

US-09-766-511B-53 (1-209) x AK014135 (1-1200)

QY	30	LeuLeuSerAlaCysPheIleValSerCysValValThrTyRHisPheThrTyrgly	48
Db	287	CTGCTGTGACAACTTATTCTCCGGTGCCTCATCATCTTGTTTACAAAATATTCTCAG	346
QY	49	-----GluthrGlyLysArgLeuSerGluLeuHisSerTyRHisSerSerLeuthrCys	66
Db	347	CTTCTTGAAGAAAAAATGATTATATAAGAACTGAATCTACTACTGGAATGGAGTGTACAAA	406
QY	67	PheSerGluGlyThrLysValProAlaThrPglCysCysProAlaSerTrpLysSerPhe	86
Db	407	TGGGCTTCACTCTTGGAGACAAAGTCTGGAGCTGTGGCCAAAGGATTGGAAGCCGTT	466
QY	87	GlySerSerCysTyRPhelIleSerSerGlu---GluLysValTrpSerLysSerGluGln	105
Db	467	GGTTCTTACTGCTACTTCCACTTCAACTGACTTGGTGGCATCTTGAATGAGAGTAAGAG	526
QY	106	AsnCysValGluMetGlyAlaHisLeuValValPheAsnThrGluAlaGluGlnAsnPhe	125
Db	527	AC-TGCTTCCACATGGGTGCTCATCTGCTGGTGATCCACAGCCAGGACACAGGATTC	585
QY	126	IleValGlnGlnLeuAsnGluSerPheSerTyRPhelGlyLeuSerAspProGlnGly	145

Percent Similarity:	57.92%	Conservative:	26
Best Local Similarity:	43.72%	Mismatches:	67
Query Match:	32.79%	Indels:	12
DB:	9	Gaps:	2

US-09-766-511B-53 (1-209) x AA868502 (1-1013)

Qy	29	AlaLeuLeuSerAlaCysPheIleValSerCysValValThrTyrHisPheThrTyrGly	48
Db	679	AGCCTCAGGACCGAGTGCCTCAACAATTTATGTATG-----	644
Qy	49	GlutThrGlyLysArgLeuSerGluLeuHisSerTyrHisSerSerLeu-----ThrCys	66
Db	643	CARAACTCCAGAGTGCCTCCAGTTACGAGGTTCACGGTCTCACTCAAGCGCACTGC	584
Qy	67	PheSerGluGlyThrLysValProAlaTrpGlyCysCysProAlaSerTrpLysSerPhe	86
Db	583	GTCTCTGAAGATAAGTACATAGAAGATTGGAGCTGCTGCCCAACCCC-TGGACTTCATTC	525
Qy	87	GlySerSerCysTyrPheIleSerSerGluGluLysValTrpSerLysSerGluGluAsn	106
Db	524	AGTCTAGTTGCTACTATATTTCTTACTGG-GATTGCATCTTGGACTAAGAGTCAAAGCAC	466
Qy	107	CysValGluMetGlyAlaHisLeuValValPheAsnThrGluAlaGluGluAsnPheIle	126
Db	465	TCCTCTGTGATGCGCTGATCTGGTGGTGATCACCCCGAGGGAACAGGATTTCACT	406
Qy	127	ValGlnGlnLeuAsnGluSerPheSerTyrPheLeuGlyLeuSerAspProGlnGlyAsn	146
Db	405	ATTCAAGAATCTGAAGAAATTTCTTCTATTATTTCTGGGCTGTCAATCCAGGGGGTCGG	346
Qy	147	AsnAsnTrpGlnTrpIleAspLysThrProTyrGluLysAsnValArgPheTrpHisLeu	166
Db	345	CGACTTTGGCAATGGGTGTACACAGACCACTACAATGAACTTCTCAATCTTGGCACTCA	286
Qy	167	GlyGluProAsnHisSerAlaGluGlnCysAlaSerIleValPheTrpLysProThrGly	186
Db	285	GGTGAACCCCAATACCTTGATGAGCGTTGGCGATATAAATTTTCCTCGTTCTTCAGACGAA	226
Qy	187	TrpGlyTrpAsnAspValIleCysGluThrArgArgAsnSerIleCysGluMetAsnLys	206
Db	225	TGGGGCTGGGAATGACATTCACCTGTCTCATGTACCTCAGAAATCAATTTGCAAGATCAAGAAG	166
Qy	207	IleTyrLeu	209
Db	165	ATCTACATC	157

RESULT 13	AI949910/c	AI949910	753 bp	mRNA	linear	EST 08-MAR-2000
LOCUS						
DEFINITION		wq04a07.x1	NCI_CGAP_Kid12	Homo sapiens	cDNA clone	IMAGE:2470260.3,
		similar to	SW:LECI_MOUSE	P24721	ASIALOGLYCOPROTEIN	RECEPTOR 2 ; ,
					mRNA	sequence.

ACCESSION	AI949910
VERSION	AI949910.1
KEYWORDS	EST.

**SOURCE ORGANISM**

REFERENCE  
1 (bases 1 to 753)  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 55)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicqap>

**TITLE**  
Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

**Tissue Procurement:** Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: M Bento Soares, Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.  
CDNA Library Arrayed by: Greg Lennon Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
 Insert length: 1065 Std Error: 0.00  
 Seq primer: -400P from Gibco  
 High quality sequence stop: 449.

FEATURES	SOURCE
----------	--------

```

1: 1723
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clones="IMAGE:2470260"
/clone_lib="NCI_CGAP_Kid12"
/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Plasmid DNA from the normalized library NCI_CGAP_Kid5 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneids 1323912-1325831, 1471368-1472903 and
1492104-1493255). Subtraction by Bento Soares and M.
Fatima Ronaldo."

```

BASE COUNT	230 a	171 c	138 g	212 t	2 others
ORIGIN					

Alignment Scores:		
Pred. No.:	7,058-33	753
Score:	377.50	67
Percent Similarity:	65.93%	Conservative: 22
Best Local Similarity:	49.63%	Mismatches: 45
Query Match:	32.40%	Indels: 1
DB:	9	Gaps: 1

US-09-766-511B-53 (1-209) x AI949910 (1-753)

Qy	76	TrpGlyCysCysProAlaSerTrpLysSerPheGlySerSerCysTyrPheIleSerSer	95
Db	732	TGGCAGTTGTGCCCAAGAATTCGNGGTCAATTAGTTCCAACCTGCATCTTATTTCCTACT	673
Qy	96	GluGluLysValTrpSerLysSerGluGlnAsnCysValGluMetGlyAlaHisLeuVal	115
Db	672	GAATCAGCATCTGGCGAAGACAGCTGGAAGAGCATGTGCTAGAAATGNAGGCTCACCTGCTGTG	613
Qy	116	ValPheAsnThrGluAlaGluGlnAsnPheIleValGlnLeuAsnGluLuserPheSer	135
Db	612	GTGATAACACTCAAGAAGACAGGAGTTTCATCTCCAGAATCTGCAAGAAGAATCTGCCT	553
Qy	136	TyrPheLeuGlyLeuSerAspProGlnGlyAsnAsnTrpGlnTrpIleAspLysThr	155
Db	552	TATTTTGTGGGGCTCTCAGATCCAGAGAGTTCAGCGACATTTGGCAATGGGTGTGATCAGACA	493
Qy	156	ProTyrGluLysAsnValArgPheTrpHisLeuGlyGluProAsnHisSerAlaGluGln	175
Db	492	CCATACAAATGAAGTTCCACATTTGGCATCTCAGGTAGCGCCAGTGCATCCCAATGAGCGC	433
Qy	176	CysAlaSerIleValPheTrpLys--proThrGlyTrpGlyTrpAsnAspValIleCys	194
Db	432	TGCGTTGTGCTAAATTTTCGTAAATCATCCCAAAAGATGGGGCTGGAATGATGTAAATTGT	373
Qy	195	GluThrArgArgAsnSerIleCysGluMetAsnLysIleTyrLeu	209
Db	372	CTTGGTCCTCAAGAGTCAGTTTGTGAGATGATGAAGATCCACCTTA	328

RESULT 14

BE650539

LOCUS

### DEFINITION

2  
C  
F  
C  
C  
E  
C  
C  
A

ACCESSION  
WEEDON

**VERSION**  
**KEYWORDS**

LOCUS	BE650539	561 bp	linear	EST 06-SEP-2000
DEFINITION	UI-M-BH2.3-aqf-f-04-0-UI.r1 NIH-BMAP_M_S3.3 Mus musculus cdna clone			
	UI-M-BH2.3-aqf-f-04-0-UI.5		nrna	
	UI-M-BH2.3-aqf-f-04-0-UI.1		nrna sequence.	

ACCESSION BE650539.1  
VERSION BE650539.1 GI:9976363

**KEYWORDS**

SOURCE house mouse.









GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 22:38:30 : Search time 36 Seconds  
(without alignments)  
773.594 Million cell updates/sec

Title: US-09-766-511b-53  
Perfect score: 1165  
Sequence: 1 MMQEQPQSTERRGWSLRL.....NDVICETRRNSICENMKIYL 209

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1165	100.0	209	AAU00479	Human TANGO 405 pr
2	885	76.0	162	AAU27446	Human SDCMP3 polyp
3	793	68.1	209	AAW63010	Mouse dectin-2
4	793	68.1	209	AAU27447	Mouse SDCMP3 polyp
5	793	68.1	209	AAU00482	Murine TANGO 405 a
6	679	58.3	167	AAW63022	Mouse dectin-2 ext
7	659	56.6	175	AAW63018	Mouse dectin-2 iso
8	637	54.7	213	AAW63018	Amino acid sequenc
9	630	54.1	218	AAE12079	Dendritic cell (DC
10	617	53.0	211	AAE12080	Dendritic cell (DC

11	592.5	50.9	168	19	AAW63020	Mouse dectin-2 iso
12	591	50.7	145	19	AAW63017	Mouse dectin-2 iso
13	521	44.7	178	22	AAU00480	Murine TANGO 405 p
14	493	42.3	134	19	AAW63019	Mouse dectin-2 iso
15	459.5	39.4	148	23	AAU98015	Human partial dend
16	459	39.4	134	23	AAU98013	Human partial dend
17	458	39.3	237	20	AAW88124	Primate DCMP1 C-le
18	458	39.3	237	23	AAW88124	Human dendritic ce
19	458	39.3	237	23	ABW90277	Human polypeptide
20	458	39.3	243	22	AAU19814	Human novel extrac
21	458	39.3	243	23	ABP48034	Human polypeptide
22	458	39.3	246	22	AAU19653	Human novel extrac
23	458	39.3	246	22	ABP47873	Human polypeptide
24	436	37.4	238	20	AAW88128	Rodent DCMP1 C-lec
25	398.5	34.2	219	20	AAU13403	Amino acid sequenc
26	398.5	34.2	219	22	AAU29035	Human PRO polypept
27	398.5	34.2	219	22	AAU74697	Human membrane ass
28	398.5	34.2	219	22	AAU80271	Human PRO244 prote
29	390.5	33.5	214	22	AAU97225	Murine C-type lect
30	334.5	28.7	85	19	AAW69237	Mouse dectin-2 His
31	252	21.6	287	18	AAU15246	Asialoglycoprotein
32	251.5	21.6	316	20	AAW88125	Primate DCMP2 C-le
33	250	21.5	273	20	AAW88129	Variant primate DC
34	250	21.5	319	23	ABG66703	Human novel polype
35	249	21.4	229	18	AAU15251	Asialoglycoprotein
36	249	21.4	270	18	AAU15252	Asialoglycoprotein
37	249	21.4	342	23	ABW95436	Human angiogenesis
38	249	21.4	342	23	ABW84830	Human secreted pro
39	249	21.4	457	20	AAW73633	Human polypeptide
40	249	21.4	542	20	AAW73634	Human secreted pro
41	249	21.4	542	23	ABP61865	Human polypeptide
42	249	21.4	542	23	ABP61865	Human polypeptide
43	249	21.4	547	22	AAU27236	Human EXMAB-14 SEQ
44	249	21.4	562	22	AAU1472	Human polypeptide
45	248	21.3	325	21	AAU94345	Human cell surface

ALIGNMENTS

RESULT 1  
AAU00479  
ID AAU00479 standard; Protein; 209 AA.

AC AAU00479;

DT 04-JUL-2001 (first entry)

XX Human TANGO 405 protein.

DE Human: TANGO 210; clone jthLa152h06; TANGO 364; TANGO 366; dectin-2;

KW INTERCEPT 394; INTERCEPT 400; TANGO 405; cellular process regulator;

KW gene therapy; growth modulator; proliferation; cell differentiation;

KW lymphocyte; bone marrow cell migration; leukaemia; lymphoma;

KW autoimmune disorder.

OS Homo sapiens.

EH Key Location/Qualifiers

FT Peptide 1..48

FT /label= Signal\_peptide

FT /note= "Specifically claimed"

FT Protein 49..209

FT /label= Mature\_TANGO\_405\_protein

FT /note= "Specifically claimed"

PN WO200118016-A1.

XX 15-MAR-2001.

PD 30-JUN-2000; 2000WO-US18174.

PF 10-SEP-1999; 99US-0393996.

PR





Db 177 VSIYVWPNPSKGNWDFCDKSHNSICEMKKIYL 209  
 |||:| | : ||||| |::: ||||| ||||

## RESULT 5

ID AAU00482 standard; Protein; 209 AA.  
 XX AAU00482;

XX 04-JUL-2001 (first entry)

XX Murine TANGO 405 alternative splice variant protein.

XX Murine; TANGO 210; clone jtmMa025a11; TANGO 364; TANGO 366; dectin-2;  
 KW INTERCEPT 400; TANGO 405; cellular process regulator;  
 KW gene therapy; growth modulator; proliferation; cell differentiation;  
 KW lymphocyte; bone marrow cell migration; leukaemia; lymphoma;  
 KW autoimmune disorder.

XX Mus sp.

XX WO200118016-A1.

XX 15-MAR-2001.

XX 30-JUN-2000; 2000WO-US18174.

XX 10-SEP-1999; 99US-0399996.

XX (MILL-) MILLENNIUM PHARM INC.

XX Fraser CC, Sharp JD, Wrighton N, Myers PS, Goodearl ADJ;

XX WPI; 2001-183280/18.

XX N-PSDB; AAS01378.

XX Isolated nucleic acid molecules encoding proteins useful as modulating  
 agents in regulating a variety of cellular processes are used for  
 treating e.g. cancer and autoimmune disorders -

XX Disclosure; Fig 6N-6P; 326pp; English.

XX The present sequence representing murine TANGO 405 alternative splice  
 variant protein is isolated from cDNA clone jtmMa025a11 from a long-term  
 bone marrow cDNA library. TANGO 405 (AAU00480) is 1 of 3 novel murine  
 proteins which include TANGO 210 (AAU00470) and INTERCEPT 400  
 (AAU00477). Six novel human proteins which include TANGO 210 (AAU00469),  
 TANGO 364 (AAU00471), TANGO 366 (AAU00472), INTERCEPT 394 (AAU00473),  
 INTERCEPT 400 (AAU00476) and TANGO 405 (AAU00479), and a rat  
 INTERCEPT 400 (AAU00478) sequence are also described. The nucleic acids  
 encoding these novel proteins are useful as modulating agents in  
 regulating a variety of cellular processes and can be used to express  
 the proteins in a host cell in gene therapy applications. Human and  
 murine TANGO 405 proteins show sequence homology to murine dectin-2.  
 TANGO 405 modulates growth, proliferation, survival, differentiation,  
 activity, morphology and movement/migration of human lymphocytes and  
 bone marrow cells and tissues and can be used to prevent, diagnose or  
 treat leukaemia, lymphomas and autoimmune disorders.

XX Sequence 209 AA;

XX Query Match 68.1%; Score 793; DB 22; Length 209;

XX Best Local Similarity 68.1%; Pred. No. 3,6e-73;

XX Matches 145; Conservative 19; Mismatches 41; Indels 8; Gaps 4;

Oy 1 MWQSQQPOSTEKRG--WLSRLMSVAGISTALLSACFVSCVTVYHFTYGTGRRLSELH 58  
 |:|:| | : | :|:| | | :|:| | | :|:| | | :|:| |

Db 1 MVQERQSQG---KGVCW-TLRLWSAAVISMLLSTCFSTASCVTVYQFTMDQFSRLYELH 56  
 |:|:| | | | | | | | | | | | | | | | | | | | | | | |

Oy 59 SYHSSLTCFSEGTGV--PANGCCPASWKSFGSGCYFISSEKVKWSKSPQNCVMAHLV 116  
 |:|:| | | | | | | | | | | | | | | | | | | | | | | |

Db 57 TYHSSLTCFSEGTWVSEKMWGCCPNHWKSGGSCYLIISTKENFNWSTSQNCVQMAHLV 116  
 |:|:| | | | | | | | | | | | | | | | | | | | | | | |

Oy 117 FNTAEQNFIVQQLNESFSYFLGLSDPQGNWQWIDTTPYKFNVRFWHLGEPNHSADQC 176  
 |:|:| | | | | | | | | | | | | | | | | | | | | | | |

Db 117 INTEAQNFITQQLNESLSYFLGLSDPQGNKQWIDTTPFSQNVRFWHPHEPNLPEERC 176  
 |:|:| | | | | | | | | | | | | | | | | | | | | | | |

Oy 177 ASIVFWKPTGWNNDVICETRRNSICEMKKIYL 209  
 |:|:| | : | | | | | :|:| | | | | | | | | | | |

Db 177 VSIYVWPNPSKGNWDFCDKSHNSICEMKKIYL 209  
 |:|:| | : | | | | | :|:| | | | | | | | | | | |

## RESULT 6

AAW63022

ID AAW63022 standard; Protein; 167 AA.

XX AAW63022;

XX 09-NOV-1998 (first entry)

XX Mouse dectin-2 extracellular domain.

XX Dectin-1; dendritic cell; c-type lectin; mouse; immunity; adjuvant;  
 KW allergy; autoimmune disease; gene therapy; vaccine; diagnosis;  
 KW drug screening.

XX Mus sp.

XX WO9828332-A2.

XX 02-JUL-1998.

XX 22-DEC-1997; 97WO-US23761.

XX 20-DEC-1996; 96US-0772440.

XX (TEXA ) UNIV TEXAS SYSTEM.

XX Arizumi K, Takashima A;

XX WPI; 1998-377594/32.

XX N-PSDB; AAW42554.

XX Nucleic acid encoding dendritic cell specific peptide(s) dectin-1  
 and -2 - useful, e.g. to regulate immune response, as vaccine  
 adjuvants, for diagnosis and drug screening

XX Disclosure; Page 153; 200pp; English.

XX This polypeptide comprises the extracellular domain of a novel  
 murine dendritic cell (DC) specific member of the C-type lectin  
 family, termed DC-associated C-type lectin-2, or dectin-2 (see also  
 AAW63010), that is essential for DC-mediated T cell activation.  
 The invention provides: dectin-1 and -2 polypeptides (see also  
 AAW63009-22 and AAW69236-37), useful for purifying T cells, for  
 detecting autoantibodies and for up-regulating immunity e.g. as  
 vaccine adjuvants; dectin DNA (see AAW42548-54, AAW42558-60 and  
 AAW44850-51); expression vectors; recombinant host cells; probes  
 and primers; antibodies; compounds that modulate dectin-mediated  
 activation of T cells; transgenic animals; and dectin ligands.

XX Dectin expression can also be down-regulated to treat allergy and  
 autoimmune disease. The extracellular domain has been expressed  
 as a His-tagged recombinant fusion protein (see AAW69237) useful for  
 raising antibodies.

XX Sequence 167 AA;

XX Query Match 58.3%; Score 679; DB 19; Length 167;

XX Best Local Similarity 70.7%; Pred. No. 1.4e-61;

XX Matches 118; Conservative 14; Mismatches 33; Indels 2; Gaps 1;

Oy 45 FTYGTGKRLSELHSHSLTCFSEGTGV--PANGCCPASWKSFGSGCYFISSEKVKWSK 102  
 |:|:| | : | | | | | | | | | | | | | | | | | | | |

Db 1 FIMQPSRRLYELHYHSLTCFSEGTWVSEKMWGCCPNHWKSGGSCYLIISTKENFNWST 60  
 |:|:| | : | | | | | | | | | | | | | | | | | | | |

```

QY 103 SEQNCVEMGAHLVYVTEAQNFIVOQLNESFSYFLGLSDPGQNNWQWIDKTPYEKNVR 162
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 SEQNCVQMGHVLVYVTEAQNFIQQLNESLSYFLGLSDPGQNGKWQWIDDTFPFSONVR 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 163 FWHLGEPNHSAEQCASTVFWKPTGNGWNDVICETRRNSICEMKKIYL 209
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 FWHPEPNLPEERCVSIVYWNPSKWGNDVFCDSKINSICEMKKIYL 167
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
AAW63018
ID AAW63018 standard; Protein; 175 AA.
XX
AC AAW63018;
XX
DT 09-NOV-1998 (first entry)
XX
DE Mouse dectin-2 isoform T-gamma.
XX
KW Dectin-2; dendritic cell; C-type lectin; mouse; immunity; adjuvant;
KW allergy; autoimmune disease; therapy; vaccine; diagnosis;
KW drug screening.
XX
OS Mus sp.
XX
PN WO9828332-A2.
XX
PD 02-JUL-1998.
XX
PF 22-DEC-1997; 97WO-US23761.
XX
PR 20-DEC-1996; 96US-0772440.
XX
PA (TEXA ) UNIV TEXAS SYSTEM.
XX
XX Ariizumi K, Takashima A;
XX WPI; 1998-377594/32.
XX
PT Nucleic acid encoding dendritic cell specific peptide(s) dectin-1
PT and -2 - useful, e.g. to regulate immune response, as vaccine
PT adjuvants, for diagnosis and drug screening
XX
PS Claim 10; Page 149; 200pp; English.
XX
XX This polypeptide comprises a truncated isoform, designated T-gamma,
CC of dectin-2 (see AAW63010), a novel murine dendritic cell (DC)
CC specific member of the C-type lectin family that is essential for
CC DC-mediated T cell activation. It contains a 34-amino acid
CC deletion primarily in the connecting domain compared with the
CC full-length (alpha) dectin-2 polypeptide. It is contemplated that
CC DC in normal tissues express mRNA for different naturally-occurring
CC isoforms of dectin-2. The invention provides: dectin-1 and
CC dectin-2 polypeptides (see also AAW63009-22 and AAW69236-37), useful
CC for purifying T cells, for detecting autoantibodies and for
CC up-regulating immunity e.g. as vaccine adjuvants; dectin DNA (see
CC AAW42548-54, AAW42558-60 and AAW44850-51); expression vectors;
CC recombinant host cells; probes and primers; antibodies; compounds
CC that modulate dectin-mediated activation of T cells; transgenic
CC animals useful for studying dectin function and for drug screening;
CC and dectin ligands. Dectin expression can be downregulated for
CC treating allergy and autoimmune disease.
XX
SQ Sequence 175 AA;
    Query Match 56.6%; Score 659; DB 19; Length 175;
    Best Local Similarity 58.3%; Pred. No. 1.6e-59;
    Matches 123; Conservative 16; Mismatches 34; Indels 38; Gaps 4;

QY 1 MMQEQQPOSTEKRG--WLSRLWSVAGISIALLSACFIVSCVVTYHFTYGETGKRLSELH 58
    |||:| | :| ||||| | | :||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 MVQERQSQG---KGVCV-TLRLWSAAVISMLLLSTCFIASCV----- 38
    |::|| | :| ||||| | | :||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

QY 59 SYHSSILTCFSEGTKVPAMGCCPASMKSCGSSCYFISSEKVKSKSEQNCVEMGAHLVYVFN 118
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 39 -----EKMGMCCCPENHKKSFSSCYLILSTKENFWSSTSEQNCVQMGHVLVYIN 84
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 119 TEAEQNFIVOQLNESFSYFLGLSDPGQNNWQWIDKTPYEKNVRFWHLGEPNHSAEQCAS 178
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 85 TEAEQNFIQQLNESLSYFLGLSDPGQNGKWQWIDDTFPFSONVRFWHPHPNLPBERCVS 144
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 179 IVFWKPTGNGWNDVICETRRNSICEMKKIYL 209
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 145 IVIWNPSKWGNDVFCDSKINSICEMKKIYL 175
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 8
AAB84215
ID AAB84215 standard; Protein; 213 AA.
XX
AC AAB84215;
XX
DT 06-AUG-2001 (first entry)
XX
DE Amino acid sequence of BDCA-2 antigen isoform.
XX
KW BDCA-2; hematopoietic cell; dendritic cell; BDCA-3; BDCA-4;
KW viral infection; autoimmune disease; allergic response; cancer.
XX
OS Homo sapiens.
XX
PN WO200136487-A2.
XX
PD 25-MAY-2001.
XX
PF 15-NOV-2000; 2000WO-IB01832.
XX
PR 15-NOV-1999; 99US-0165555.
PR 23-NOV-1999; 99US-0167076.
PR 28-JAN-2000; 2000US-0179003.
PR 07-FEB-2000; 2000US-0180775.
PR 11-APR-2000; 2000US-0196824.
PR 13-APR-2000; 2000US-0197205.
XX
PA (MILT-) MILTENYI BIOTECH GMBH.
XX
PI Schmitz J, Dzionek A, Buck DW;
XX
XX WPI; 2001-355622/37.
DR N-PSDB; AAF90241.
XX
XX Compositions and cell populations enriched in dendritic cells through
PT use of antigen-binding fragments specific for BDCA-2, BDCA-3 or
PT BDCA-4, are used to treat viral infections, autoimmune diseases,
PT allergic responses and cancer -
XX
PS Claim 122; Fig 5; 115pp; English.
XX
XX The present sequence represents an antigen designated BDCA-2. The
CC specification describes compositions and hematopoietic cell populations
CC enriched in dendritic cells (DCs). These compositions are produced using
CC antigen-binding fragments specific for BDCA-2, BDCA-3 or BDCA-4. The DCs
CC obtained are used to treat viral infections, autoimmune disease, allergic
CC response, and cancer. BDCA-1, BDCA-2, BDCA-3 and BDCA-4 monoclonal
CC antibodies and their antigen-binding fragments are used to detect,
CC enumerate and isolate DC populations from leukapheresis material, whole
CC blood and tonsils and from non-hematopoietic and hematopoietic tissues.
XX
SQ Sequence 213 AA;
    Query Match 54.7%; Score 637; DB 22; Length 213;
    Best Local Similarity 54.0%; Pred. No. 3.8e-57;
    Matches 115; Conservative 30; Mismatches 64; Indels 4; Gaps 2;

QY 1 MMQEQQPOSTEK-RGWLRLWSVAGISIALLSACFIVSCVVTYHFTYGETGKRLSEL-- 57
    |::|| | | |::||| :|| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 1 MVBEPQDREKGLWFWQLKVMWMAVVSILLSVCTTVSSVPHNFMVSKTVKRLSKLRE 60

QY 58 -HSYHSLTFCFSEGTKVPAGCCPASWKSFGSGCYFISSEKVKWSEKQNCVEMGAHLV 116

Db 61 YQYHPSLTCTVMEGKDIEDWSCCPTPTWTFQSSCYFISTGMSQSWTKSOKNCSVMGADLV 120

QY 117 FNTAEQNFIVQOLNESFSYFLGLSDPQGNNNWQIDKTPYBKNVRFWHLGEPNHSABQC 176

Db 121 INTREQDFTIQNLKRNSYFLGLSDPGGRHWQVDDTPYNNVTFWHSGEPPNLDERC 180

QY 177 ASIVFNKPTGWNVDVICTRRNSICEMNKIYL 209

Db 181 AINFRSSEEWGNDIHCHVPOKSICKMKKIYI 213

RESULT 9

AAE12079

ID AAE12079 standard; Protein; 218 AA.

XX

AC AAE12079;

XX

XX 18-DEC-2001 (first entry)

XX

DE Dendritic cell (DC) DCLEC protein.

XX

KW Dendritic cell; DC; DCLEC protein; gene therapy; dermatological; vaccine;

KW atopic dermatitis; autoimmune disease; inflammatory skin disease; cancer;

KW immunosuppressive; AIDS; Acquired immune deficiency syndrome; cytostatic;

KW chromosomal identification; pharmaceutical; hypersensitivity; virucide;

KW transplant rejection; chronic inflammatory disease; anti-HIV.

XX

OS Unidentified.

XX

XX Key Location/Qualifiers

FX Domain 26..46

FT /note= "Single transmembrane region"

FT Domain 114..211

FT /note= "Extracellular C-type lectin domain"

XX

XX WO200172773-A2.

XX

XX 04-OCT-2001.

XX

XX 28-MAR-2001; 2001WO-EP03542.

XX

XX 29-MAR-2000; 2000US-192934P.

PR 18-MAY-2000; 2000US-205026P.

PR 18-MAY-2000; 2000US-205026P.

PR 19-MAY-2000; 2000US-205767P.

PR 19-MAY-2000; 2000US-205769P.

XX

XX (NOVS ) NOVARTIS AG.

PA (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.

PA Werner G, Phares W, Jaritz M, Lapp H, Kalthoff FS;

XX

XX WPI; 2001-616466/71.

DR N-PSDB; AAD19729.

XX

XX New polypeptides for screening therapeutic agonists and antagonists

PT comprise dendritic cell polypeptides -

XX

XX Claim 2; Page 51; 52pp; English.

XX

XX The invention relates to dendritic cell (DC) proteins and their

CC corresponding DNA molecules. A pharmaceutical composition comprising

CC agonist and antagonist of DC proteins are useful for treating abnormal

CC conditions related to both an excess of and insufficient level of

CC expression of DC gene, or related to both an excess of and insufficient

CC activity of DC protein. Soluble form of DC proteins are used as an active

CC ingredient in combination with pharmaceutical acceptable carriers.

CC DC genes and proteins are useful for treating chronic inflammatory

CC diseases, autoimmune diseases, transplant rejection crisis, including

CC inflammatory skin diseases such as contact hypersensitivity, atopic

CC dermatitis or virally-induced immune suppression such as AIDS and cancer.

CC DC protein is useful for inducing immunological response in a mammal, and

CC as immunogen to produce antibodies immunospecific for the polypeptide.

CC DC gene is useful in gene therapy. DC gene is also useful as a diagnostic

CC reagent, and for chromosomal identification. The present sequence is

CC of dendritic cell (DC) DCLEC protein which is found to belong to the family

CC of C-type lectins with one single carbohydrate recognition domain at the

XX C-terminal end.

XX

SQ Sequence 218 AA;

Query Match 54.1%; Score 630; DB 22; Length 218;

Best Local Similarity 54.5%; Pred. No. 2.1e-56;

Matches 114; Conservative 28; Mismatches 63; Indels 4; Gaps 2;

QY 5 QOQOSTEK-RGWLRLWSVAGISIALLSACFIVSCVVVTHFYTGTKRLSEL---HSY 60

Db 10 QEPQDREKGLWFWQLKVMWMAVVSILLSVCTTVSSVPHNFMVSKTVKRLSKLREYQY 69

QY 61 HSSLTFCFSEGTKVPAGCCPASWKSFGSGCYFISSEKVKWSEKQNCVEMGAHLVFNTE 120

Db 70 HPSLTCTVMEGKDIEDWSCCPTPTWTFQSSCYFISTGMSQSWTKSOKNCSVMGADLVNTR 129

QY 121 AEQNFIVQOLNESFSYFLGLSDPQGNNNWQIDKTPYBKNVRFWHLGEPNHSABQCASIV 180

Db 130 EQQDFIQLNKRNSYFLGLSDPGGRHWQVDDTPYNNVTFWHSGEPPNLDERCALIN 189

QY 181 FWKPTGWNVDVICTRRNSICEMNKIYL 209

Db 190 FRSEEWGNDIHCHVPOKSICKMKKIYI 218

RESULT 10

AAE12080

ID AAE12080 standard; Protein; 211 AA.

XX

AC AAE12080;

XX

XX 18-DEC-2001 (first entry)

XX

DE Dendritic cell (DC) DCLEC/SPLICE 1 variant protein.

XX

KW Dendritic cell; DC; DCLEC protein; gene therapy; dermatological; vaccine;

KW atopic dermatitis; autoimmune disease; inflammatory skin disease; cancer;

KW immunosuppressive; AIDS; Acquired immune deficiency syndrome; cytostatic;

KW chromosomal identification; pharmaceutical; hypersensitivity; virucide;

KW transplant rejection; chronic inflammatory disease; anti-HIV; variant.

XX

OS Unidentified.

XX

XX Key Location/Qualifiers

FX Misc-difference 1 /note= "Encoded by C"

FT

XX WO200172773-A2.

XX

XX 04-OCT-2001.

XX

XX 28-MAR-2001; 2001WO-EP03542.

XX

XX 29-MAR-2000; 2000US-192934P.

PR 18-MAY-2000; 2000US-205026P.

PR 18-MAY-2000; 2000US-205026P.

PR 19-MAY-2000; 2000US-205767P.

PR 19-MAY-2000; 2000US-205769P.

XX

XX (NOVS ) NOVARTIS AG.

PA (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.

XX Werner G, Phares W, Jaritz M, Lapp H, Kalthoff FS;

XX

XX WPI; 2001-616466/71.

DR

DR N-PSDB; AAD19730.

XX New polypeptides for screening therapeutic agonists and antagonists

PT comprise dendritic cell polypeptides -

XX

PS Claim 2; Page 52; 52pp; English.

XX

CC The invention relates to dendritic cell (DC) proteins and their

CC corresponding DNA molecules. A pharmaceutical composition comprising

CC agonist and antagonist of DC proteins are useful for treating abnormal

CC conditions related to both an excess of and insufficient level of

CC expression of DC gene, or related to both an excess of and insufficient

CC activity of DC protein. Soluble form of DC proteins are used as an active

CC ingredient in combination with pharmaceutical acceptable carriers.

CC DC genes and proteins are useful for treating chronic inflammatory

CC diseases, autoimmune diseases, transplant rejection crisis, including

CC inflammatory skin diseases such as contact hypersensitivity, atopic

CC dermatitis or virally-induced immune suppression such as AIDS and cancer.

CC DC protein is useful for inducing immunological response in a mammal, and

CC as immunogen to produce antibodies immunospecific for the polypeptide.

CC DC gene is useful in gene therapy. DC gene is also useful as a diagnostic

CC reagent, and for chromosomal identification. The present sequence is

CC dendritic cell (DC) DCLEC/SPLICE 1 variant protein which is found to

CC belong to the family of C-type lectins with one single carbohydrate

CC recognition domain at the C-terminal end.

XX

SQ Sequence 211 AA;

Query Match 53.0%; Score 617; DB 22; Length 211;

Best Local Similarity 52.9%; Pred. No. 4.3e-55;

Matches 111; Conservative 30; Mismatches 63; Indels 6; Gaps 2;

QY 6 OPOSTEKRR---GWLRLRLWSVAGISIALLSACFTVSCVTVYHFTYGTGKRLSEL---HS 59

DB 2 RPAEDREKGLWFWQLKVMNAVVISLLLSVCFVTSSVPHNFMYSKTVRLSKLREYQQ 61

QY 60 YHSSLTCTFSECTKVPAMGCCPASWKSFGSSCYFTISSBEKVMKSEQNCVEMGAHLVVNT 119

DB 62 YHPSLTCVMEGKDIEDWSCCPTPTWTSFQSSCYFTISGMSQKSWKSNKCVMGADLVVINT 121

QY 120 EAEONFIVQOLNESFSYFLGLSDPGQNNNQWIDKTPYEKNVRFWHLGEPNHSAEQCAI 179

DB 122 REEQDFITONLKRNSYFLGLSDPGRRRHQWVDQTPYNENVTFWHSGEPNLDERCAII 181

QY 180 VFWRPTGWNVDVICETRRNSICEMNKIYL 209

DB 182 NFRSEEWGWNDIHCHVPQKSIKMKKIYI 211

RESULT 11

AAW63020

ID AAW63020 standard; Protein; 168 AA.

XX

AC AAW63020;

XX

DT 09-NOV-1998 (first entry)

XX

DE Mouse dectin-2 isoform T-epsilon.

XX

KW Dectin-2; dendritic cell; c-type lectin; mouse; immunity; adjuvant;

KW allergy; autoimmune disease; therapy; vaccine; diagnosis;

KW drug screening.

XX

OS Mus sp.

XX

PN WO9828332-A2.

XX

PD 02-JUL-1998.

XX

PF 22-DEC-1997; 97WO-US23761.

XX

PR 20-DEC-1996; 96US-0772440.

XX

PA (TEXA ) UNIV TEXAS SYSTEM.

XX

PI Ariizumi K, Takashima A;

XX

DR WPI; 1998-377594/32.

XX

PT Nucleic acid encoding dendritic cell specific peptide(s) dectin-1

PT and -2 - useful, e.g. to regulate immune response, as vaccine

PT adjuvants, for diagnosis and drug screening

XX

PS Claim 12; Page 151; 200pp; English.

XX

CC This polypeptide comprises a truncated isoform, designated

CC T-epsilon, of dectin-2 (see AAW63010), a novel murine dendritic cell

CC (DC) specific member of the C-type lectin family that is essential

CC for DC-mediated T cell activation. T-epsilon contains a 43-amino

CC amino acid deletion in the carbohydrate recognition domain compared

CC with the full-length (alpha) dectin-2 polypeptide. It is

CC contemplated that DC in normal tissues express mRNA for different

CC naturally-occurring isoforms of dectin-2. The invention provides:

CC dectin-1 and dectin-2 polypeptides (see also AAW63009-22 and

CC AAW69236-37), useful for purifying T cells, for detecting

CC autoantibodies and for up-regulating immunity e.g. as vaccine

CC adjuvants; dectin DNA (see AAW42548-54, AAW42558-60 and AAW44850-51);

CC expression vectors; recombinant host cells; probes and primers;

CC antibodies; compounds that modulate dectin-mediated activation of T

CC cells; transgenic animals useful for studying dectin function and

CC for drug screening; and dectin ligands. Dectin expression can be

CC downregulated for treating allergy and autoimmune disease.

XX

SQ Sequence 168 AA;

Query Match 50.9%; Score 592.5; DB 19; Length 168;

Best Local Similarity 55.4%; Pred. No. 1e-52;

Matches 118; Conservative 16; Mismatches 30; Indels 49; Gaps 5;

QY 1 MMQEQOQPQSTEKRG--WLSRLRLWSVAGISIALLSACFTVSCVTVYHFTYGTGKRLSELH 58

DB 1 MVQERSQSG--KGVCW-TLRLWSAAVISMLLLSTCFIASCVVTYQFINDQPSRRLYELH 56

QY 59 SYHSSLTCTFSECTKV--PAWGCCPASWKSFGSSCYFTISSBEKVMKSEQNCVEMGAHLVV 116

DB 57 TVHSSLTCTFSECTVMSEKMWGCCPNHWKSGSSCYLSTKENFWSSTSEQNCVQMGHLVV 116

QY 117 FNTAEQNFIVQOLNESFSYFLGLSDPGQNNNQWIDKTPYEKNVRFWHLGEPNHSAEQC 176

DB 117 INTEAEQNFITQOLNESLSYFLGLSN----- 142

QY 177 ASIVFWKPTGWNVDVICETRRNSICEMNKIYL 209

DB 143 -----PSKWGWNVDVFCDSKHSICEMKKIYL 168

RESULT 12

AAW63017

ID AAW63017 standard; Protein; 145 AA.

XX

AC AAW63017;

XX

DT 09-NOV-1998 (first entry)

XX

DE Mouse dectin-2 isoform T-beta.

XX

KW Dectin-2; dendritic cell; c-type lectin; mouse; immunity; adjuvant;

KW allergy; autoimmune disease; therapy; vaccine; diagnosis;

KW drug screening.

XX

OS Mus sp.

XX

PN WO9828332-A2.

XX

PD 02-JUL-1998.

XX

PF 22-DEC-1997; 97WO-US23761.  
PR 20-DEC-1996; 96US-0772440.  
XX (TEXA ) UNIV TEXAS SYSTEM.  
PA Arizumi K, Takashima A;  
XX WPI; 1998-377594/32.  
XX Nucleic acid encoding dendritic cell specific peptide(s) dectin-1  
PT and -2 - useful, e.g. to regulate immune response, as vaccine  
PT adjuvants, for diagnosis and drug screening  
XX  
PS Claim 9; Page 148-149; 200pp; English.  
XX  
XX This polypeptide comprises a truncated isoform, designated T-beta,  
CC of dectin-2 (see AAW63010), a novel murine dendritic cell (DC)  
CC specific member of the C-type lectin family that is essential for  
CC DC-mediated T cell activation. It contains a 65-amino acid  
CC deletion primarily in the transmembrane and connecting domains  
CC compared with the full-length (alpha) dectin-2 polypeptide. It is  
CC contemplated that DC in normal tissues express mRNA for different  
CC naturally-occurring isoforms of dectin-2. The invention provides:  
CC dectin-1 and dectin-2 polypeptides (see also AAW63009-22 and  
CC AAW69236-37), useful for purifying T cells, for detecting  
CC autoantibodies and for up-regulating immunity e.g. as vaccine  
CC adjuvants; dectin DNA (see AAV42548-54, AAV42558-60 and AAV44850-51);  
CC expression vectors; recombinant host cells; probes and primers;  
CC antibodies; compounds that modulate dectin-mediated activation of  
CC T cells; transgenic animals useful for studying dectin function  
CC and for drug screening; and dectin ligands. Dectin expression can  
CC be downregulated for treating allergy and autoimmune disease.  
XX  
SQ Sequence 145 AA;  
  
Query Match 50.7%; Score 591; DB 19; Length 145;  
Best Local Similarity 73.9%; Pred. No. 1.2e-52;  
Matches 99; Conservative 11; Mismatches 24; Indels 0; Gaps 0;  
  
QY 76 WGCCPASWKSFGSGCYFISSEKRWKSRQNCVEMGAHLVFTAEQNFIVQQLNESFS 135  
DB 12 WGCCPNHWKSGSGCYLISTKFNFWSTSEQNCVOMGAHLVINTAEQNFITQQLNESLS 71  
  
QY 136 YFLGLSDPGNKNWOWIDTPEYKKNRVFWHLGEPNHSABQACASIVFKPTGKWNDVICE 195  
DB 72 YFLGLSDPGQNGKQWIDDTPTSEQNVRFWHPHPNLPPEERCVSIVVWNPFGKGNWDFCD 131  
  
QY 196 TRNSICEKMKIYL 209  
DB 132 SKHNSICEKMKIYL 145  
  
RESULT 13  
AAU00480  
ID AAU00480 standard; Protein; 178 AA.  
XX AAU00480;  
XX  
XX 04-JUL-2001 (first entry)  
XX Murine TANGO 405 protein.  
DE  
XX  
XX Murine; TANGO 210; clone jtmMa025a11; TANGO 364; TANGO 366; dectin-2;  
KW INTERCEPT 394; INTERCEPT 400; TANGO 405; cellular process regulator;  
KW gene therapy; growth modulator; proliferation; cell differentiation;  
KW lymphocyte; bone marrow cell migration; leukaemia; lymphoma;  
KW autoimmune disorder.  
OS  
XX Mus sp.  
XX  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..42

FT  
FT Protein /label= Signal\_peptide  
FT 43..178  
XX /label= Mature\_TANGO\_405\_protein  
XX  
XX WO200118016-A1.  
XX  
XX 15-MAR-2001.  
XX  
XX 30-JUN-2000; 2000WO-US18174.  
XX  
XX 10-SEP-1999; 99US-0393996.  
XX  
XX (MILL-) MILLENNIUM PHARM INC.  
XX  
XX Fraser CC, Sharp JD, Wrighton N, Myers PS, Goodearl ADJ;  
XX  
XX WPI; 2001-183280/18.  
XX N-PSDB; AAS01376.  
XX  
XX Isolated nucleic acid molecules encoding proteins useful as modulating  
PT agents in regulating a variety of cellular processes are used for  
PT treating e.g. cancer and autoimmune disorders -  
XX  
XX Disclosure; Fig 6E-6F; 326pp; English.  
XX  
XX The present sequence representing murine TANGO 405 protein is isolated  
CC from cDNA clone jtmMa025a11 from a long-term bone marrow cDNA library.  
CC An alternative splice variant of this protein is also described  
CC (AAU00482). Other novel murine proteins include TANGO 210 (AAU00470) and  
CC INTERCEPT 400 (AAU00477). Six novel human proteins which include  
CC TANGO 210 (AAU00469), TANGO 364 (AAU00471), TANGO 366 (AAU00472),  
CC INTERCEPT 394 (AAU00473), INTERCEPT 400 (AAU00476) and TANGO 405  
CC (AAU00479), and a rat INTERCEPT 400 (AAU00478) sequence are also  
CC described. The nucleic acids encoding these novel proteins are useful as  
CC modulating agents in regulating a variety of cellular processes and can  
CC be used to express the proteins in a host cell in gene therapy  
CC applications. Human and murine TANGO 405 proteins show sequence homology  
CC to murine dectin-2. TANGO 405 modulates growth, proliferation, survival,  
CC differentiation, activity, morphology and movement/migration of human  
CC lymphocytes and bone marrow cells and tissues and can be used to prevent,  
CC diagnose or treat leukaemia, lymphomas and autoimmune disorders.  
XX  
SQ Sequence 178 AA;  
  
Query Match 44.7%; Score 521; DB 22; Length 178;  
Best Local Similarity 66.5%; Pred. No. 2.4e-45;  
Matches 105; Conservative 12; Mismatches 31; Indels 10; Gaps 5;  
  
QY 1 MMQEQQPOSTEKRG--WLSURLNSVAGISTALLSACFIVSCVVTYHYTGETKRLSELH 58  
DB 1 MVQERQSQG---KGVCW-TLRWSAAVISMILLSTCTFIASCVVTYQFIMQPSRRLYELH 56  
  
QY 59 SYHSSLTCFSGTKV--PANGCCPASWKSFGSGCYFTSSSEKRWKSRQNCVEMGAHLV 116  
DB 57 TYHSSLTCFSGTKVSWSEKMMGCCPNHWKSGSGCYLISTKFNFWSTSEQNCVOMGAHLV 116  
  
QY 117 FNTEAEQNFIVQQLNESFSYFLGLSDPG--GNNNNWOWI 152  
DB 117 INTEAEQNFITQQLNESLSYFLGLSDPKVMANGNGSMI 154  
  
RESULT 14  
AAW63019  
ID AAW63019 standard; Protein; 134 AA.  
XX  
XX AAW63019;  
XX  
XX 09-NOV-1998 (first entry)  
XX  
XX Mouse dectin-2 isoform T-delta.  
XX  
XX Dectin-2; dendritic cell; c-type lectin; mouse; immunity; adjuvant;  
KW allergy; autoimmune disease; therapy; vaccine; diagnosis;



## drug screening.

KW XX OS Mus sp.  
 XX PN WO9828332-A2.  
 XX PD 02-JUL-1998.  
 XX PF 22-DEC-1997; 97WO-US23761.  
 XX PR 20-DEC-1996; 96US-0772440.  
 XX PA (TEXA ) UNIV TEXAS SYSTEM.  
 XX PI Ariizumi K, Takashima A;  
 XX DR WPI; 1998-377594/32.  
 XX  
 PT Nucleic acid encoding dendritic cell specific peptide(s) dectin-1  
 PT and -2 - useful, e.g. to regulate immune response, as vaccine  
 PT adjuvants, for diagnosis and drug screening  
 XX  
 PS Claim 10; Page 149-150; 200pp; English.  
 XX  
 CC This polypeptide comprises a truncated isoform, designated T-delta,  
 CC of dectin-2 (see AAW63010), a novel murine dendritic cell (DC)  
 CC specific member of the C-type lectin family that is essential for  
 CC DC-mediated T cell activation. T-delta contains a 41-amino acid  
 CC deletion in the C-terminus of the carbohydrate recognition domain  
 CC as well as a 34-amino acid deletion primarily in the connecting  
 CC domain compared with the full-length (alpha) dectin-2 polypeptide.  
 CC It is contemplated that DC in normal tissues express mRNA for  
 CC different naturally-occurring isoforms of dectin-2. The invention  
 CC provides: dectin-1 and dectin-2 polypeptides (see also AAW63009-22  
 CC and AAW6236-37), useful for purifying T cells, for detecting  
 CC autoantibodies and for up-regulating immunity e.g. as vaccine  
 CC adjuvants; dectin DNA (see AAW42548-54, AAW42558-60 and AAW44850-51);  
 CC expression vectors; recombinant host cells; probes and primers;  
 CC antibodies; compounds that modulate dectin-mediated activation of T  
 CC cells; transgenic animals useful for studying dectin function and  
 CC for drug screening; and dectin ligands. Dectin expression can be  
 CC downregulated for treating allergy and autoimmune disease.  
 XX  
 SO Sequence 134 AA;  
 Query Match 42.3%; Score 493; DB 19; Length 134;  
 Best Local Similarity 57.5%; Pred. No. 1.3e-42;  
 Matches 96; Conservative 10; Mismatches 23; Indels 38; Gaps 4;  
 Qy 1 MMQEQQPOSTEKRG--WLSRLMSVAGISIALLSACFIVSCVVTYHFTYGETGKRILSELH 58  
 Db 1 MVOERQSQG---KGVGW-TLRLWSAAVISMILLSTCFIASCV----- 38  
 Qy 59 SYHSLTCFSECTKVPWAGCCPASPWSKSGSCYFISSEKVKSEONCVEMGAHLVVPN 118  
 Db 39 -----EKWVGCCPNHWKSGSCYLIISTKENFWSTSEONCVQMG AHLV IN 84  
 Qy 119 TEAQNFITVQOLNESFSYFLGLSDPQGNWIDKTPYEKNVREWH 165  
 Db 85 TEAQNFITVQOLNESFSYFLGLSDPQGNWIDKTPYEKNVREWH 131  
 RESULT 15  
 AAU98015  
 ID AAU98015 standard; Protein; 148 AA.  
 XX  
 AC AAU98015;  
 XX  
 DT 27-AUG-2002 (first entry)  
 XX  
 DE Human partial dendritic cell immunoreceptor #2.  
 XX  
 KW Human; dendritic cell immunoreceptor; cytostatic;

KW antiasthmatic; anorectic; antidiabetic; cancer; allergy; anaphylaxis;  
 KW asthma; inflammation; Obesity; diabetes; central nervous system disorder;  
 KW Alzheimer's disease; Parkinson's disease; dementia; osteoarthritis;  
 KW cardiovascular disorder; myocardial infarction; ischaemic heart disease;  
 KW congestive heart failure; chronic obstructive pulmonary disease; COPD.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200232958-A2.  
 XX  
 PD 25-APR-2002.  
 XX  
 PF 12-OCT-2001; 2001WO-BP11812.  
 XX  
 PR 16-OCT-2000; 2000US-24096P.  
 PR 27-AUG-2001; 2001US-314661P.  
 XX  
 PA (FARB ) BAYER AG.  
 XX  
 PI Smolyar A;  
 XX  
 DR WPI; 2002-463308/49.  
 DR N-PSDB; ABX52914.  
 XX  
 PT Novel human dendritic cell immunoreceptor polypeptide, useful for  
 PT treating cancer, asthma, obesity, diabetes, central nervous system  
 PT disorder or cardiovascular disorder -  
 XX  
 PS Claim 25; Fig 16; 114pp; English.  
 XX  
 CC The invention relates to a purified human dendritic cell immunoreceptor.  
 CC polypeptide and its encoding nucleic acid. Also included are an  
 CC expression vector comprising the nucleic acid, a host cell containing  
 CC the vector, reducing the activity of human dendritic cell immunoreceptor  
 CC involves contacting a cell with a reagent which specifically binds to  
 CC of the polypeptide or nucleic acid and a reagent that modulates the activity  
 CC of the polypeptide or nucleic acid. The reagent is useful for the  
 CC preparation of a medicament for modulating the activity of human  
 CC dendritic cell immunoreceptor in a disease such as cancer, allergy,  
 CC anaphylaxis, asthma, inflammation, obesity, diabetes, a central nervous  
 CC system (CNS) disorder (e.g. Alzheimer's disease, Parkinson's disease or  
 CC dementia), or a cardiovascular disorder (e.g. myocardial infarction,  
 CC ischaemic heart disease, congestive heart failure), chronic obstructive  
 CC pulmonary disease (COPD) and osteoarthritis (many other diseases and  
 CC disorders are listed in the specification). The polypeptide and nucleic  
 CC acid are useful for identifying test compounds which act as agonists or  
 CC antagonists, for raising specific antibodies, and as a bait protein in a  
 CC two-hybrid or three-hybrid assay. The nucleic acid is useful in  
 CC diagnostic assays for detecting diseases and abnormalities or  
 CC susceptibility to disease and abnormalities related to the presence of  
 CC mutations. The present sequence represents a partial sequence of a  
 CC dendritic cell immunoreceptor.

Sequence 148 AA;

Query Match 39.4%; Score 459.5; DB 23; Length 148;  
 Best Local Similarity 54.1%; Pred. No. 3.9e-39;  
 Matches 80; Conservative 20; Mismatches 47; Indels 1; Gaps 1;

Qy 63 SUTCFSEGTKV-PAMGCCPASPWSKSGSCYFISSEKVKSEONCVEMGAHLVVFNTFA 121  
 Db 1 S VNISSSLIWEWDSCTPTWTSFQSSCYFISTGMQSWTKSKNCSVMGADLVVINTRE 60  
 Qy 122 EQNFIVQOLNESFSYFLGLSDPQGNWIDKTPYEKNVREWHGEPNHSSEQASIVF 181  
 Db 61 EQDFIIONLKRNSYFLGLSDPGRHRWQWVDQTPYENVTFWHSGEPNLDCAIINF 120  
 Qy 182 WKPTGWGNDVICETERNISCEMNNIYL 209  
 Db 121 RSSEWGWNDIHCVPQKSIKMKKIYI 148

Search completed: February 11, 2003, 22:56:57

Job time : 37 secs



```
QY 59 SYHSSLTCSFSEGTKV--PAWGCPASWKSFGSSCYFISSEKVKWSKSEQNCVEMGAHLVV 116
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 57 YHSSLTCSFSEGMTVSEKMWGCCPNHWSFGSSCYFISSEKVKWSKSEQNCVEMGAHLVV 116
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 117 FYTEARQNFIVQOLNESFSYFLGLSDPOGNNWQIDKTPYEKNVRFWHLGEPNHSABQC 176
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 117 INTEARQNFITQOLNESLSYFLGLSDPOGNGKWQIDDTFFSQNVRFVHPHPNLPBRC 176
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 177 ASIVFKPTGWNVDVICETRNSENCKIYL 209
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 177 VSIYWNPSKWGNDFCDSKHNSICEMKKIYL 209
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
RESULT 2
US-08-772-440-21
; Sequence 21, Application US/08772440
; Patent No. 6046158
; GENERAL INFORMATION:
; APPLICANT: Ariizumi, Kiyoshi
; APPLICANT: Takashima, Akira
; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE
; TITLE OF INVENTION: LECTINS, DECTIN-1 AND DECTIN-2; COMPOSITIONS AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTXD:493
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 167 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-772-440-21
Query Match 58.3%; Score 679; DB 3; Length 167;
Best Local Similarity 70.7%; Pred. No. 6.7e-65;
Matches 118; Conservative 14; Mismatches 33; Indels 2; Gaps 1;
QY 45 FTYGTEGKRLSELHSHSLTCSFSEGTKV--PAWGCPASWKSFGSSCYFISSEKVKWSK 102
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 FIMDQPSRRLYELHTYHSLTCSFSEGMTVSEKMWGCCPNHWSFGSSCYFISSEKVKWSK 60
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 103 SEQNCVEMGAHLVVFVTEARQNFIVQOLNESFSYFLGLSDPOGNNWQIDKTPYEKNVR 162
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 SEQNCVEMGAHLVVFVTEARQNFITQOLNESLSYFLGLSDPOGNGKWQIDDTFFSQNVR 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 163 FWHLGPENHSABQCAISIVFKPTGWNVDVICETRNSENCKIYL 209
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 FWHPHPNLPBRCVSIYWNPSKWGNDFCDSKHNSICEMKKIYL 167
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
RESULT 3
```

```
US-08-772-440-15
; Sequence 15, Application US/08772440
; Patent No. 6046158
; GENERAL INFORMATION:
; APPLICANT: Ariizumi, Kiyoshi
; APPLICANT: Takashima, Akira
; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE
; TITLE OF INVENTION: LECTINS, DECTIN-1 AND DECTIN-2; COMPOSITIONS AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTXD:493
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 175 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-772-440-15
Query Match 56.8%; Score 659; DB 3; Length 175;
Best Local Similarity 58.3%; Pred. No. 9.7e-63;
Matches 123; Conservative 16; Mismatches 34; Indels 38; Gaps 4;
QY 1 MMEQQQPQSTSEKRG--WLSRLWSVAGISTALLSACFIVSCVVTYHFTYGETCKRLSELH 58
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MVQERQSQG---KGVCM-TLRLWSAAVISMLLLSTCFIASCV----- 38
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 59 SYHSSLTCSFSEGTKVPAWGCPASWKSFGSSCYFISSEKVKWSKSEQNCVEMGAHLVVF 118
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 39 -----EKMWGCCPNHWSFGSSCYFISSEKVKWSKSEQNCVEMGAHLVVF 84
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 119 TEARQNFIVQOLNESFSYFLGLSDPOGNNWQIDKTPYEKNVRFWHLGEPNHSABQCA 178
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 85 TEARQNFITQOLNESLSYFLGLSDPOGNGKWQIDDTFFSQNVRFVHPHPNLPBRCVS 144
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 179 IVFKPTGWNVDVICETRNSENCKIYL 209
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 145 IVYWNPSKWGNDFCDSKHNSICEMKKIYL 175
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
RESULT 4
US-08-772-440-17
; Sequence 17, Application US/08772440
; Patent No. 6046158
; GENERAL INFORMATION:
; APPLICANT: Ariizumi, Kiyoshi
; APPLICANT: Takashima, Akira
; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE
; TITLE OF INVENTION: LECTINS, DECTIN-1 AND DECTIN-2; COMPOSITIONS AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 42
```

```
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/772,440
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTXD:493
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 168 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
;
; US-08-772-440-17
```

```
Query Match 50.9%; Score 592.5; DB 3; Length 168;
Best Local Similarity 55.4%; Pred. No. 1.2e-55;
Matches 118; Conservative 16; Mismatches 30; Indels 49; Gaps 5;

QY 1 MMQBQPOQTEKRG--WLSURLNSVAGISIALLSACFIYSCVYVYHYFTYGETKRLSEH 58
:||||| : : ||||| :||| ||||| : : ||||| :
Db 1 MVQBRSQG---KGVCV-TURLNSAAVISMLLSTCFIASCVVYQFINDQPSRLLYELH 56

QY 59 SYHSLTCFSEGRKV--PANGCCPASWKSFGSSCYFISSEKVKWSKQNCVEMGAHLVV 116
:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 57 TYHSLTCFSEGRKVMVSEKMGWCCPNHMKFSGSSCYLSTKFNFWSTSEQNCVQMGHLVV 116
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 117 FNTAEONFIVQOLNESFSYFLGLSDPDGNNNQWIDKTPYEKNVRFWHLGEPNHSAEQC 176
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 117 INTEAEONFTQOLNESLSFLGLSN----- 142

QY 177 ASIVFWKPTGWNNDVICETRRNSICEMNKIYL 209
:||||| :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 143 -----PSKWGNDVFCDSKHNSICEMKKIYL 168
```

```
RESULT 5
US-08-772-440-14
; Sequence 14, Application US/08772440
; Patent No. 6046158
; GENERAL INFORMATION:
; APPLICANT: Ariizumi, Kiyoshi
; APPLICANT: Takashima, Akira
; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE
; TITLE OF INVENTION: LECTINS, DECTIN-1 AND DECTIN-2; COMPOSITIONS AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/772,440
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTXD:493
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 23:
; COMPUTER: IBM PC compatible
```

```
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/772,440
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTXD:493
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
;
; US-08-772-440-14

Query Match 50.7%; Score 591; DB 3; Length 145;
Best Local Similarity 73.9%; Pred. No. 1.4e-55;
Matches 99; Conservative 11; Mismatches 24; Indels 0; Gaps 0;

QY 76 WGCCPASWKSFGSSCYFISSEKVKWSKQNCVEMGAHLVVFNTAEONFIVQOLNESFS 135
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 12 WGCCPNHMKFSGSSCYLSTKFNFWSTSEQNCVQMGHLVVFNTAEONFTQOLNESLS 71

QY 136 YFLGLSDPDGNNNQWIDKTPYEKNVRFWHLGEPNHSAEQCASIVFWKPTGWNNDVICE 195
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 72 YFLGLSDPDGNGKQWIDDPFSQNVREHPHPNLPPEKRCVSIYVWPNKSGWNDVFC 131

QY 196 TRNSICEMNKIYL 209
:||||| |||||

Db 132 SKHNSICEMKKIYL 145

RESULT 6
US-08-772-440-23
; Sequence 23, Application US/08772440
; Patent No. 6046158
; GENERAL INFORMATION:
; APPLICANT: Ariizumi, Kiyoshi
; APPLICANT: Takashima, Akira
; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE
; TITLE OF INVENTION: LECTINS, DECTIN-1 AND DECTIN-2; COMPOSITIONS AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/772,440
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTXD:493
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 23:
; COMPUTER: IBM PC compatible
```

```
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 131 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-772-440-23

Query Match          48.5%; Score 565; DB 3; Length 131;
Best Local Similarity 73.3%; Pred. No. 7.2e-53;
Matches 96; Conservative 11; Mismatches 24; Indels 0; Gaps 0;

QY 79 CPASWKSFGSSCYFISSEKVKWSQNCVEMGAHLVFNTEAEQNFIQQQLNESFSYFL 138
   || ||||| |||: || |||||: ||||| ||||| ||||| ||||| |||||
Db 1 CPNHWSFGSSCYLISTKENFWSTSEQNCVQMGHLVFNTEAEQNFIQQQLNESLSYFL 60

QY 139 GLSDPGQNNWQWIDKTPYEKNVRFWHLGEPNHSAEQCAIVFWKPTGQWNDVICETR 198
   ||||| ||||| |||: ||||| ||| ||||| ||| ||||| ||| ||||| |||:
Db 61 GLSDPGQNGKWQWIDTTPFSQNVRFWHPHEPPLPERCVSIYVWNPCKGWNDFCDSKH 120

QY 199 NSICEMNKIYL 209
   ||||| ||||
Db 121 NSICEMKKIYL 131

RESULT 8
US-08-772-440-16
; Sequence 16, Application US/08772440
; Patent No. 6046158
; GENERAL INFORMATION:
; APPLICANT: Ariizumi, Kiyoshi
; APPLICANT: Takashima, Akira
; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE
; TITLE OF INVENTION: LECTINS, DECTIN-1 AND DECTIN-2; COMPOSITIONS AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/772,440
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTXD:493
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 134 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-772-440-16

Query Match          42.3%; Score 493; DB 3; Length 134;
Best Local Similarity 57.5%; Pred. No. 3.6e-45;
Matches 96; Conservative 10; Mismatches 23; Indels 38; Gaps 4;

QY 1 MMQEQPQOSTEKRG--WLSLRLWSVAGISTALLSACFIVCVVTHFTYGETGKRLSELH 58
   |||: ||| |||: ||| ||||| |||: ||||| ||| ||||| |||
Db 1 MVQERQSQG---KGVCH-TLRLWSAAVISMILLSTCFIASCV----- 38

QY 59 SYHSSLTCFSEGTKVPWAGCCPASWKSFGSSCYFISSEKVKWSQNCVEMGAHLVFN 118
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 39 -----EKMWGCCPNHWSFGSSCYLISTKENFWSTSEQNCVQMGHLVVIN 84

QY 119 TEAEQNFIQQQLNESFSYFLGLSDPGQNNWQWIDKTPYEKNVRFW 165
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 85 TEAEQNFIQQQLNESLSYFLGLSDPGQNGKWQWIDTTPFSQNVRFW 131

RESULT 9
US-09-111-470-2
; Sequence 2, Application US/09111470
; Patent No. 6277959
```

```
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 131 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-772-440-23

Query Match          48.5%; Score 565; DB 3; Length 131;
Best Local Similarity 73.3%; Pred. No. 7.2e-53;
Matches 96; Conservative 11; Mismatches 24; Indels 0; Gaps 0;

QY 79 CPASWKSFGSSCYFISSEKVKWSQNCVEMGAHLVFNTEAEQNFIQQQLNESFSYFL 138
   || ||||| |||: || |||||: ||||| ||||| ||||| ||||| |||||
Db 1 CPNHWSFGSSCYLISTKENFWSTSEQNCVQMGHLVFNTEAEQNFIQQQLNESLSYFL 60

QY 139 GLSDPGQNNWQWIDKTPYEKNVRFWHLGEPNHSAEQCAIVFWKPTGQWNDVICETR 198
   ||||| ||||| |||: ||||| ||| ||||| ||| ||||| ||| ||||| |||:
Db 61 GLSDPGQNGKWQWIDTTPFSQNVRFWHPHEPPLPERCVSIYVWNPCKGWNDFCDSKH 120

QY 199 NSICEMNKIYL 209
   ||||| ||||
Db 121 NSICEMKKIYL 131

RESULT 7
US-08-772-440-27
; Sequence 27, Application US/08772440
; Patent No. 6046158
; GENERAL INFORMATION:
; APPLICANT: Ariizumi, Kiyoshi
; APPLICANT: Takashima, Akira
; TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE
; TITLE OF INVENTION: LECTINS, DECTIN-1 AND DECTIN-2; COMPOSITIONS AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/772,440
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTXD:493
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 131 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-772-440-27

Query Match          48.5%; Score 565; DB 3; Length 131;
Best Local Similarity 73.3%; Pred. No. 7.2e-53;
Matches 96; Conservative 11; Mismatches 24; Indels 0; Gaps 0;

QY 79 CPASWKSFGSSCYFISSEKVKWSQNCVEMGAHLVFNTEAEQNFIQQQLNESFSYFL 138
   || ||||| |||: || |||||: ||||| ||||| ||||| ||||| |||||
Db 1 CPNHWSFGSSCYLISTKENFWSTSEQNCVQMGHLVFNTEAEQNFIQQQLNESLSYFL 60
```







```
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (650)852-9196
;   TELEX:      (650)496-1200
;   INFORMATION FOR SEQ ID NO: 10:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 273 amino acids
;       TYPE:    amino acid
;       TOPOLOGY: linear
;     MOLECULE TYPE: protein
US-09-111-470-10

Query Match               21.5%; Score 250; DB 4; Length 273;
Best Local Similarity     32.4%; Pred. No. 8.7e-19;
Matches 55; Conservative 28; Mismatches 55; Indels 32; Gaps

QY 53 RLSELHSHSSLTG-----FSEGTKVPAMGCCPCASMKSFSGSCSYFLSSEEKVM 100
    |:| |||| |::| ||| :| ||||: |
DB 106 RVQLVLDLKLTCOVATLNNGEEASTEGT-----CCPVNVEHQDSFYFSHGMSW 159
    ||:| | ||||: | ||||: | ||||: |
QY 101 KSKEQNCVENGAHLVFNTAEQNFIVOQLNESFSYFLGLSDPQGNNNMOWIDKTPTYEKN 160
    :::|| | ||||: | ||||: | ::|||: | ||||: |
DB 160 AEAKEYCOLNAHLIVNSREEQNFVKYLGSAYTW-MGLSDPEG--AKKWVDGTDYATG 216
    :||:| | ||||: | ||||: | ||||: |
QY 161 VRFWHLGEPNH-----SAEQCASIFVFKPTGMGNNDVICETFRNSICE 203
    :|:|:| | || |||:| | |||:|:| :|||
DB 217 FQNKPGQPDDWGHLGGGEDCA---HFHPDG-RWNDDVCQRPHYHWCE 262
    :||:| | || |||:| | |||:|:| :|||

RESULT 15
US-08-688-342-4
; Sequence 4, Application US/08688342
; Patent No. 5871964
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Goli, Surya K.
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: NOVEL HUMAN C-TYPE LECTIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/688,342
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0095-1 CIP
TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 292 amino acids
;   TYPE:    amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1235724
US-08-688-342-4
```

Query Match 21.5%; Score 250; DB 2; Length 292;  
Best Local Similarity 32.4%; Pred. No. 9.5e-19;  
Matches 55; Conservative 28; Mismatches 55; Indels 32; Gaps 7;  
QY 53 RLSELHSHSLTC-----FSEGTKVPAMGCCPASWKSFGSSCYFISSEK YW 100  
Db 125 RVQQLVQDLKKLTCQVATLNNNGEEASTEGT-----CCPVNWVEHQDSCYWFESHGMSW 178  
QY 101 SKSEQNCVEMGAHLVVFNTEAEQNFIVQQLNFSFSYFLGLSDPQGNNNWQWIDKTPYEKN 160  
Db 179 AEAEKYCQLKNAHLVVINSREEQNFQKYLGSAYTW-MGLSDPEG--AKKWVDGTDYATG 235  
QY 161 VRFWHLGEPNH-----SAEQCASIVFWKPTGCGWNDVICETRNSICE 203  
Db 236 FQWKFQCPDDWQGHGLGGEDCA---HFHPDG-RWNDDVCQRPYHWVCE 281

Search completed: February 11, 2003, 22:59:48  
Job time : 15 secs

GenCore version 5.1.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 22:59:51 : Search time 11 Seconds  
(without alignments)  
485.429 Million cell updates/sec

Title: US-09-766-511B-53  
Perfect score: 1165  
Sequence: 1 MMQEQQPQSTKRGWLSRL.....NDVICETRRNSICEMNKIYL 209

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 140259 seqs, 25548876 residues

Total number of hits satisfying chosen parameters: 140259

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA.\*  
1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pcp.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pcp.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pcp.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pcp.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pcp.\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pcp.\*  
7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pcp.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pcp.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pcp.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pcp.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pcp.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pcp.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pcp.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	637	54.7	213	12	US-10-090-466-2
2	535.5	46.0	182	12	US-10-090-466-4
3	458	39.3	237	9	US-09-862-802-2
4	458	39.3	237	9	US-09-870-759-49
5	458	39.3	243	10	US-09-764-870-464
6	458	39.3	246	10	US-09-764-870-303
7	436	37.4	238	9	US-09-862-802-8
8	398.5	34.2	219	9	US-09-905-291A-377
9	398.5	34.2	219	9	US-09-965-529-3
10	398.5	34.2	219	9	US-09-902-853-377
11	398.5	34.2	219	9	US-09-907-824-377
12	398.5	34.2	219	9	US-09-907-841-377
13	398.5	34.2	219	9	US-09-904-011-377
14	398.5	34.2	219	9	US-10-174-590-24
15	398.5	34.2	219	9	US-10-176-758-24
16	398.5	34.2	219	9	US-10-175-737-24
17	398.5	34.2	219	9	US-09-906-742-377
18	398.5	34.2	219	9	US-10-173-706-24
19	398.5	34.2	219	9	US-10-175-738-24

20	398.5	34.2	219	9	US-10-175-752-24	Sequence 24, Appl
21	398.5	34.2	219	9	US-10-176-482-24	Sequence 24, Appl
22	398.5	34.2	219	9	US-10-176-757-24	Sequence 24, Appl
23	398.5	34.2	219	9	US-10-176-913-24	Sequence 24, Appl
24	398.5	34.2	219	9	US-10-180-552-24	Sequence 24, Appl
25	398.5	34.2	219	9	US-10-180-557-24	Sequence 24, Appl
26	398.5	34.2	219	9	US-09-906-838-377	Sequence 24, Appl
27	398.5	34.2	219	9	US-09-907-613-377	Sequence 377, App
28	398.5	34.2	219	9	US-09-907-942-377	Sequence 377, App
29	398.5	34.2	219	9	US-10-173-700-24	Sequence 24, Appl
30	398.5	34.2	219	9	US-10-174-572-24	Sequence 24, Appl
31	398.5	34.2	219	9	US-10-174-579-24	Sequence 24, Appl
32	398.5	34.2	219	9	US-10-174-582-24	Sequence 24, Appl
33	398.5	34.2	219	9	US-10-174-588-24	Sequence 24, Appl
34	398.5	34.2	219	9	US-10-175-739-24	Sequence 24, Appl
35	398.5	34.2	219	9	US-10-175-740-24	Sequence 24, Appl
36	398.5	34.2	219	9	US-10-175-743-24	Sequence 24, Appl
37	398.5	34.2	219	9	US-10-176-488-24	Sequence 24, Appl
38	398.5	34.2	219	9	US-10-176-492-24	Sequence 24, Appl
39	398.5	34.2	219	9	US-10-176-747-24	Sequence 24, Appl
40	398.5	34.2	219	9	US-10-176-750-24	Sequence 24, Appl
41	398.5	34.2	219	9	US-10-176-985-24	Sequence 24, Appl
42	398.5	34.2	219	9	US-10-176-987-24	Sequence 24, Appl
43	398.5	34.2	219	9	US-10-176-991-24	Sequence 24, Appl
44	398.5	34.2	219	9	US-10-176-992-24	Sequence 24, Appl
45	398.5	34.2	219	9	US-10-176-993-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1  
US-10-090-466-2  
; Sequence 2, Application US/10090466  
; Patent No. US20020137914A1  
; GENERAL INFORMATION:  
; APPLICANT: Turner, C. Alexander Jr.  
; APPLICANT: Mathur, Brian  
; APPLICANT: Cullinan, Emily B.  
; TITLE OF INVENTION: No. US20020137914A1 Human Dectin Proteins and Polynucleotide  
; FILE REFERENCE: LEX-0315-USA  
; CURRENT APPLICATION NUMBER: US/10/090.466  
; CURRENT FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: US 60/274,961  
; PRIOR FILING DATE: 2001-03-12  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 213  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-10-090-466-2

Query Match	54.7%	Score 637;	DB 12;	Length 213;
Best Local Similarity	54.0%	Pred. No. 1.2e-56;		
Matches	115;	Conservative 30;	Mismatches 64;	Indels 4; Gaps 2;
Qy	1	MMQEQQPQSTK-RGWLSRLWSVAGISALLSACFTVSCVYVYHFTYGETGKRLSEL--	57	
Db	1	MVPEEPQDREKGLWMFQKWSMAVVSILLSCFTVSSVVPHPNFMYSKTVKRLSKLRE	60	
Qy	58	-HSVHSSLTCTFSEGTGKVPAMGCCPASWKSFGSSCYFISSEKVMKSEQNCVEMGAHLV	116	
Db	61	YQYHPSLTCVMEGKDIEDSCCPTPTWTSQSSCYFISTGMQSWTKSQKNCVMGADLV	120	
Qy	117	FNTAEQNFIVQOLNESFSYFLGLSDPQGNMNMWIDKTPYKRVNRFHLCGEPNHSAEQC	176	
Db	121	INTREODFIQNLKRNSYFLGLSDPGRRHMQWVDQTPYNNVTFWHSGEFNNLDERC	180	
Qy	177	ASIVFWKPTGWNVDVICETRRNSICEMNKIYL	209	
Db	181	AIINFRSSEWGNNDTHCVFPQKSICKMKKIYI	213	

```
RESULT 2
US-10-090-466-4
; Sequence 4, Application US/10090466
; Patent No. US20020137914A1
; GENERAL INFORMATION:
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Mathur, Brian
; APPLICANT: Cullinan, Emily B.
; TITLE OF INVENTION: NO. US20020137914A1 Human Dectin Proteins and Polynucleotides
; FILE REFERENCE: LEX-0315-USA
; CURRENT APPLICATION NUMBER: US/10/090.466
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 60/274,961
; PRIOR FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 182
; TYPE: PRT
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(182)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-090-466-4

Query Match 46.0%; Score 535.5; DB 12; Length 182;
Best Local Similarity 54.7%; Pred. No. 1.5e-46;
Matches 94; Conservative 23; Mismatches 52; Indels 3; Gaps 1;

Qy 41 VTYHTYGTGKRLSEL---HSYHSLTCTFSGTKVPAGWCCPASWSKSGSCYFISSEE 97
Db 11 VPHNFWYKTVARLSKRLREYQYHXSLTCVMEGKDIEDWSCCPTPTWTSFQSSCYFISIGM 70
Qy 98 KVMKSEQNCVPMGAHLVVFNTAEQNFIVQOLNESFSYFLGLSDPQNNWQIDTKTPY 157
Db 71 QSWTKSOKNCVAGDLVVFNTAEQNFIVQOLNESFSYFLGLSDPQNNWQIDTKTPY 130
Qy 158 EXNVRFHGLGEPNHSAEQCAIVFWKPTGKRWGNDVICETRRNSICEMNKIYL 209
Db 131 NENVTFHSGEPNLDRCALINFRSSEEGWNDIHCHPQKSIKMKIYI 182

RESULT 3
US-09-862-802-2
; Sequence 2, Application US/09862802
; Patent No. US20020165346A1
; GENERAL INFORMATION:
; APPLICANT: Schering-plough Corporation
; TITLE OF INVENTION: MAMMALIAN MEMBRANE PROTEIN GENES; RELATED REAGENTS
; FILE REFERENCE: SF0695B
; CURRENT APPLICATION NUMBER: US/09/862,802
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 09/111,470
; PRIOR FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: mammalian protein
US-09-862-802-2

Query Match 39.3%; Score 458; DB 9; Length 237;
Best Local Similarity 44.4%; Pred. No. 1.3e-38;
Matches 83; Conservative 34; Mismatches 66; Indels 4; Gaps 3;

Qy 26 ISTALLSACFIVSCVVTYHTYGTGKRLSELHSHYSSLTCTFSEGTGV--PAWGCCPASW 83
Db 52 IFFLLAISFFIAFVFFQ-KYSQLEKTKTKELVHTTLECVKKNMPVEETAWSCCPKNW 110
Qy 84 KSGSSCYFISSEBEKVMKSEQNCVPMGAHLVVFNTAEQNFIVQOLNESFSYFLGLSDP 143
Db 111 KFSNNCYFISTESASWQDSEKDCARMEALLVINTQEEQDFIFQNLQESAFYVGLSDP 170
Qy 144 QGNNWQIDTKPYEKNVRFHGLGEPNHSAEQCAIVFWK-PTGKRWGNDVICETRRNSIC 202
Db 171 EGRHWQWVDQTPYNESSTFWHPREPSDNPNERCVLNFRRKSPRKGWNDVNCILGPQRSVC 230
Qy 203 EMNKIYL 209
Db 231 EMMKIHL 237

RESULT 4
US-09-870-759-49
; Sequence 49, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 49
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-870-759-49

Query Match 39.3%; Score 458; DB 9; Length 237;
Best Local Similarity 44.4%; Pred. No. 1.3e-38;
Matches 83; Conservative 34; Mismatches 66; Indels 4; Gaps 3;

Qy 26 ISTALLSACFIVSCVVTYHTYGTGKRLSELHSHYSSLTCTFSEGTGV--PAWGCCPASW 83
Db 52 IFFLLAISFFIAFVFFQ-KYSQLEKTKTKELVHTTLECVKKNMPVEETAWSCCPKNW 110
Qy 84 KSGSSCYFISSEBEKVMKSEQNCVPMGAHLVVFNTAEQNFIVQOLNESFSYFLGLSDP 143
Db 111 KFSNNCYFISTESASWQDSEKDCARMEALLVINTQEEQDFIFQNLQESAFYVGLSDP 170
Qy 144 QGNNWQIDTKPYEKNVRFHGLGEPNHSAEQCAIVFWK-PTGKRWGNDVICETRRNSIC 202
Db 171 EGRHWQWVDQTPYNESSTFWHPREPSDNPNERCVLNFRRKSPRKGWNDVNCILGPQRSVC 230
Qy 203 EMNKIYL 209
Db 231 EMMKIHL 237

RESULT 5
US-09-764-870-464
; Sequence 464, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PYZ14
; CURRENT APPLICATION NUMBER: US/09/764,870
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 464
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-870-464
```



Wed Feb 19 09:55:49 2003

us-09-766-511b-53.rspt

Page 7

Job time : 30 secs

```

      : | | : : | | | | : | | | | : | | | | | | | | | |
Db 88 NELNCTKSVSPMEDKVMSCCPKDWRLFGSHCYLVPTVSSASWNKSEENCSRMGAHLVVI 147

Qy 118 NTEAEQNFIVQOLNESFSYFLGLSDPGQGNWNNWIDKTPYEKNVRFVHGLGEPNHSAQCA 177
      : : : | | | | : | | | | | : | | | | | | | | | |
Db 148 QSOEQDITGILDTHAAVFIGLMD-TGHRQWQVDDQTPYEESTFWHNGEPSSGNEKCA 206
      : : : | | | | | : | | | | | : | | | | | | | | | |

Qy 178 SIVF-WKPTGWNVDVICETRRNSICEMNKIYL 209
      : | : | | | | | : | : | : | | | | |
Db 207 TIIRWK-TGWNWDISCLQKSVQCMKKINL 238

RESULT 13
Q9H229 PRELIMINARY; PRT; 204 AA.
AC Q9H229;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE C-type lectin DBB27 short form.
GN DBB27.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Richard M., Beaulieu A.D.;
RT "A short ORF for the C-type lectin DBB27.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF200738; AAG35593.1; -.
DR InterPro; IPR002353; AntifreezeZell.
DR Pfam; PF00059; Lectin_C.1.
DR PRINTS; PR00356; ANTIFREEZEII.
DR SMART; SM00034; CLECT.1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW Lectin.
SQ SEQUENCE 204 AA; 23550 MW; 4DFFC27549CEAA5 CRC64;

Query Match 37.2%; Score 433.5; DB 4; Length 204;
Best Local Similarity 53.7%; Pred. No. 3.3e-36;
Matches 73; Conservative 22; Mismatches 40; Indels 1; Gaps 1;

Qy 75 AWGCCPASWKSFGSCYFISSEKVKSEONCEVEMGAHLVFNTEAEQNFIVQOLNESF 134
      | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 69 AWSCCPKNWKFSFSSNCYFISTESASWQSEKDCARMEAHLLVINTQEEQDFIQNJOES 128
      : | | | | | | | | | | | | | | | | | | | | | | | |

Qy 135 SYFLGLSDPGQGNWNNWIDKTPYEKNVRFVHGLGEPNHSAQCAIVFWK-PTGWNWDVI 193
      : | | | | | | | | | | | | | | | | | | | | | | | |
Db 129 AYFVGLSDPEGQRHWQVDDQTPYNESTFWHPREPSDPNERCVVLNFKSPKRWGNDVN 188
      : | | | | | | | | | | | | | | | | | | | | | | | |

Qy 194 CETRRNSICEMNKIYL 209
      : | : | | | | | : | : | : | | | | |
Db 189 CLGFQSVCEMKNKHL 204

RESULT 14
Q923C7 PRELIMINARY; PRT; 262 AA.
AC Q923C7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Similar to C-type (calcium dependent, carbohydrate recognition
DE domain) lectin, superfamily member 6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
```

```

RA Strausberg R.;
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC006623; AAH06623.1; -.
DR InterPro; IPR001304; Lectin_C.1.
DR Pfam; PF00059; lectin_c.1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
SQ SEQUENCE 262 AA; 29822 MW; 413DAD69511071B8 CRC64;

Query Match 36.9%; Score 430; DB 11; Length 262;
Best Local Similarity 37.2%; Pred. No. 1e-35;
Matches 89; Conservative 40; Mismatches 72; Indels 38; Gaps 8;

Qy 3 QEOQPOSTEKRGWLSURLNSVAGISIALLSACFIVSCVVTYHFTYGE----- 49
      : | : | | | | : | : | | | | : | : | | | | : | : |
Db 30 REAPIRDLRKPGPSULLTSLM-LLLLLLAITFLVAFII-YFOKYSOLLEEKKAANNIMH 87
      : | : | | | | | | | | | | | | | | | | | | | | | |

Qy 50 -----TGKRLSELHSYHSLTFCFSGTKVPWAGCGCPASWKSFGSSCYFI 93
      : | : | | | | | | | | | | | | | | | | | | | | | |
Db 88 NELNCTKSVSPMEAPPFIGORALTLESIEDLGLAPEDKV--WSCCPKDWRLFGSHCYLV 145
      : | : | | | | | | | | | | | | | | | | | | | | | |

Qy 94 S--SEEKVSKSEONCEVEMGAHLVFNTEAEQNFIVQOLNESFSYFLGLSDPGQNNWOW 151
      : | : | | | | | | | | | | | | | | | | | | | | | |
Db 146 PTVSSASWNKSEENCSRMGAHLVVIQSOEQDFITGITLTHAAVFIGLMD-TGHRQWOW 204
      : | : | | | | | | | | | | | | | | | | | | | | | |

Qy 152 IDKTPYEKNVRFVHGLGEPNHSAQCAIVF-WKPTGWNWDVICETRRNSICEMNKIYL 209
      : | : | | | | | | | | | | | | | | | | | | | | | |
Db 205 VDQTPYEESTFWHNGEPSSGNEKCATIIRWK-TGWNWDISCLQKSVQCMKKINL 262
      : | : | | | | | | | | | | | | | | | | | | | | | |

RESULT 15
Q8WXW9 PRELIMINARY; PRT; 230 AA.
AC Q8WXW9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Fc-epsilon receptor III.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang W., Zhang J., Li N., Wan T., Chen T., Cao X.;
RT "Novel human Fc-epsilon receptor III.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF328684; AAL56016.1; -.
DR InterPro; IPR001304; Lectin_C.1.
DR Pfam; PF00059; lectin_c.1.
DR SMART; SM00034; CLECT.1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW Receptor.
SQ SEQUENCE 230 AA; 26603 MW; 03CAB887D71BB084 CRC64;

Query Match 36.5%; Score 425; DB 4; Length 230;
Best Local Similarity 44.2%; Pred. No. 2.8e-35;
Matches 76; Conservative 31; Mismatches 61; Indels 4; Gaps 3;

Qy 26 ISIALLSACFIVSCVVTYHFTYGTGKRLSELHSYHSLTFCFSGTKV--PAMGCCPASW 83
      : | : | | | | : | : | | | | : | : | | | | : | : |
Db 52 IFFLLAISFFIAFVFFQ-KYSOLLEKTTKRELVTHTLECVKKNPVBETAMSCCPKNW 110
      : | : | | | | | | | | | | | | | | | | | | | | | |

Qy 84 KSGSSCYFISSEKVKSEONCEVEMGAHLVFNTEAEQNFIVQOLNESFSYFLGLSDP 143
      : | | | | | | | | | | | | | | | | | | | | | | | |
Db 111 KFSSSNCYFISTESASWQSEKDCARMEAHLLVINTQEEQDFIQNJOESAYFVGLSDP 170
      : | | | | | | | | | | | | | | | | | | | | | | | |

Qy 144 OGNNNWQIDKTPYEKNVRFVHGLGEPNHSAQCAIVFWK-PTGWNWDVIC 194
      : | : | | | | | | | | | | | | | | | | | | | | | |
Db 171 EQRHWQVDDQTPYNESTFWHPREPSDPNERCVVLNFKSPKRWGNDVNC 222
      : | : | | | | | | | | | | | | | | | | | | | | | |

Search completed: February 11, 2003, 23:00:24
```

```
Db 171 EGQRHWQVWDQTPYNESSTFWHPREPSDPNRCVVLNFRKSPKRWGNDVNCGLGQRSVC 230
Qy 203 BMNKIYL 209
Db 231 EMMKIHL 237

RESULT 10
Q922H6 Q922H6 PRELIMINARY; PRT; 219 AA.
AC Q922H6;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE C-type lectin.
GN CLECSF8 OR MCL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 1
RP SEQUENCE FROM N.A.
RX MEDLINE=98325084; PubMed=9660840;
RA Balch S.G., McKnight A.J., Seldin M.F., Gordon S.;
RT "Cloning of a novel C-type lectin expressed by murine macrophages.";
RL J. Biol. Chem. 273:18656-18664(1998).
DR EMBL; AF061272; AAD05125.1; .
DR HSP; P22897; IEGG.
DR MGD; MGI:1298389; Clecsf8.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW Lectin.
SO SEQUENCE 219 AA; 25619 MW; 1801A65F4B69BC40 CRC64;

Query Match 38.0%; Score 443; DB 11; Length 219;
Best Local Similarity 42.7%; Pred. No. 3.8e-37;
Matches 79; Conservative 31; Mismatches 71; Indels 4; Gaps 1;

Qy 24 AGISIALLSACFIVSCVYHYFTYGETGKRLSELHSHSLTFCSE----GTKVPWAGCC 79
| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 24 AVWSISFLSACFISCTCLVTHHYFLRWTRGSYVKLSYDHYTRVTCITREPPQPGGTGTC 83
| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Qy 80 PASMKSGSSCYTSSEKVKSEQNCVEMGAHLVVFNTAEQNEITVQOLNESFSYFLG 139
| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 84 PVSWRATFQNSCYFLPNDQTHWSESNCSGSHLVNTINTEAEQNFVTQLDKRFSTFLG 143
| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Qy 140 LSDPQGNNGNWIDKTPYEKNVRFHLGEPNHSAEQCAIVFWKPTGWNNDVICETRRN 199
| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 144 LADENVEGQWQVKTFFNPHTVFWKESGNDPWEEDCVVLVHVHEKVVWVDFPCHEVR 203
| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Qy 200 SICEM 204
| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 204 RICKL 208

RESULT 11
Q8WX18 Q8WX18 PRELIMINARY; PRT; 215 AA.
AC Q8WX18;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE C-type lectin-like receptor CLEC-6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN 1
RP SEQUENCE FROM N.A.
RA Colonna M.;
RT "C-type lectin-like receptor.";
```

```
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF411850; AAL37713.1; -.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW Receptor; Lectin.
SQ SEQUENCE 215 AA; 24674 MW; D782B5AFCEAC0354 CRC64;

Query Match 37.9%; Score 441; DB 4; Length 215;
Best Local Similarity 40.9%; Pred. No. 6e-37;
Matches 85; Conservative 40; Mismatches 73; Indels 10; Gaps 6;

Qy 5 QOQOSTEKRGLSLRLWSV-AGISIALLSACFIVSCVYHY-FTYGETGKRLSELHSHS 62
| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 4 EKPSQKLEGGMHPOLIPSVIAVVFILLGVCFIASCLVTHHNFSCRKRGTVGHL-EHHA 62
| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Qy 63 SLTFCSEGTKVP- ----WGCCPASWKSFGSSCYFISSEKVKWSKSEQNCVEMGAHLVVF 117
| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 63 KLKCIKESKSELKSABGSTWNCPCIDWRAFQSNQCYFPLTDNKTWASERNCSGMGAHLMTI 122
| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Qy 118 NTEAEQNFIVQOLNESFSYFLGSLDPQGNNGNWIDKTPYEKNVRFHLGEPNHS-ABQC 176
| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 123 STEAEQNFITQIDRLRLSYFLGLRDENAKGWRWVDQTFPNFRVFWHKNPDNSQGENC 182
| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Qy 177 ASIVFWKPTGWNNDVICETRRNSICEM 204
| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 183 VVLVY-NQDKWANDVPCNFEASRICKI 209
| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

RESULT 12
Q9QZ15 Q9QZ15 PRELIMINARY; PRT; 238 AA.
AC Q9QZ15;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Dendritic cell immunoreceptor.
GN CLECSF6 OR DCIR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 1
RP SEQUENCE FROM N.A.
RX TISSUE=LYMPH NODE;
RX MEDLINE=99370001; PubMed=10438934;
RA Bates E.E., Fournier N., Garcia E., Valladeau J., Durand I., Pin J.J.,
RA Zurawski S.M., Patel S., Abrams J.S., Lebecque S., Garrone P.,
RA Saeland S.;
RT "APCs express DCIR, a novel c-type lectin surface receptor containing
an immunoreceptor tyrosine-based inhibitory motif.";
RL J. Immunol. 163:1973-1983(1999).
DR EMBL; AJ133533; CAB57870.1; -.
DR HSP; P06734; IHLJ.
DR MGD; MGI:1349412; Clecsf6.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW Receptor.
SQ SEQUENCE 238 AA; 27323 MW; 2DB8AE0E11B18A56 CRC64;

Query Match 37.4%; Score 436; DB 11; Length 238;
Best Local Similarity 39.9%; Pred. No. 2.2e-36;
Matches 85; Conservative 45; Mismatches 73; Indels 10; Gaps 8;

Qy 3 QEOQPOSTEKRGLSLRLWSVAGISIALLSACFIVSCVYHYFTYGE-TGKRLSELHSHV 61
| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 30 REKPIRDLRPGSPSLTLTSLM-LLLLLLAIITFLVAFII-YEQKYSOLLEKKAANKIMH 87
| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Qy 62 SSLTFCSEGTKV--PAWGCPASWKSFGSSCYFIS--SEEKVKWSKSEQNCVEMGAHLVVF 117
```



```
Db 87 SOEBQDFITGILDIHAAYFIGLWD-TGHRQWQWVDQTPYEESVTFWHNGEPSSDNEKCVT 145
QY 179 IVFWKPTGWMNDVICETRRNSICEMNKIYL 209
RL : : : ||||| : : : | : | : | : |
Db 146 VYRRNIGWGWNDISCNLKQKSVQCMKINL 176
RESULT 7
QY 179 IVFWKPTGWMNDVICETRRNSICEMNKIYL 209
ID Q9UI34 PRELIMINARY; PRT; 237 AA.
AC Q9UI34;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE C-type lectin superfamily 6.
GN CLECSF6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Richard M., Beaulieu A.D.;
RT "A novel C-type lectin expressed in GM-CSF stimulated neutrophils.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF109146; AAF14348.1; -.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW Lectin.
SQ SEQUENCE 237 AA; 27488 MW; 1A68BCE8323345BA CRC64;

Query Match 39.3%; Score 458; DB 4; Length 237;
Best Local Similarity 44.4%; Pred. No. 1.2e-38;
Matches 83; Conservative 34; Mismatches 66; Indels 4; Gaps 3;

QY 26 ISIALLSACFIVSCVVTYHFTYGETGKRLSELHSYHSLTCFSEGTKV--PAWGCCPASW 83
| : : ||||| : : : | : | : | : |
Db 52 IFELLALISFFIAFVIFQ-KYSQLEKKTTELVTLECVKKNMPVEETAWSCCPKNW 110
| : : ||||| : : : | : | : | : |

QY 84 KFGSSCYFISSEKVKSEQNCVEMGAHLVVFNTAEQNFIVQOLNESFSYFLGLSDP 143
| : : ||||| : : : | : | : | : |
Db 111 KFSNNCYFISTESASQWQSEKDCARMEAHLLVINTQEQDFIFQNLQEESEAYFVGLSDP 170
| : : ||||| : : : | : | : | : |

QY 144 QGNNNWQIDKTPYEKNVRFHGLGEPNHSACASIVFWK-PTGWNNDVICETRRNSIC 202
| : : ||||| : : : | : | : | : |
Db 171 EGQRHWQWVDQTPYNESSTFWHPRESDPNRCVVLNFRKSPKRWGNDVNCGLGPQRSVC 230
| : : ||||| : : : | : | : | : |

QY 203 EMNKIYL 209
| : : |||||
Db 231 EMMKIHL 237

RESULT 8
QY 203 EMNKIYL 209
ID Q9UMR7 PRELIMINARY; PRT; 237 AA.
AC Q9UMR7;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE Dendritic cell immunoreceptor.
GN DCIR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bates E.E., Fournier N., Garcia E., Valladeau J., Durand I., Pin J.J.,
RA Zurawski S.M., Patel S., Abrams J.S., Lebecque S., Garrone P.,
```

```
RA Saeland S.;
RT "APCs express DCIR, a novel c-type lectin surface receptor containing
an immunoreceptor tyrosine-based inhibitory motif.";
RL J. Immunol. 163:1973-1983(1999).
DR EMBL; AJ133532; CAB54001.1; -.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
KW Receptor.
SQ SEQUENCE 237 AA; 27512 MW; 1D07003E4C9CF96E CRC64;

Query Match 39.3%; Score 458; DB 4; Length 237;
Best Local Similarity 44.4%; Pred. No. 1.2e-38;
Matches 83; Conservative 34; Mismatches 66; Indels 4; Gaps 3;

QY 26 ISIALLSACFIVSCVVTYHFTYGETGKRLSELHSYHSLTCFSEGTKV--PAWGCCPASW 83
| : : ||||| : : : | : | : | : |
Db 52 IFELLALISFFIAFVIFQ-KYSQLEKKTTELVTLECVKKNMPVEETAWSCCPKNW 110
| : : ||||| : : : | : | : | : |

QY 84 KFGSSCYFISSEKVKSEQNCVEMGAHLVVFNTAEQNFIVQOLNESFSYFLGLSDP 143
| : : ||||| : : : | : | : | : |
Db 111 KFSNNCYFISTESASQWQSEKDCARMEAHLLVINTQEQDFIFQNLQEESEAYFVGLSDP 170
| : : ||||| : : : | : | : | : |

QY 144 QGNNNWQIDKTPYEKNVRFHGLGEPNHSACASIVFWK-PTGWNNDVICETRRNSIC 202
| : : ||||| : : : | : | : | : |
Db 171 EGQRHWQWVDQTPYNESSTFWHPRESDPNRCVVLNFRKSPKRWGNDVNCGLGPQRSVC 230
| : : ||||| : : : | : | : | : |

QY 203 EMNKIYL 209
| : : |||||
Db 231 EMMKIHL 237

RESULT 9
QY 203 EMNKIYL 209
ID Q9NS33 PRELIMINARY; PRT; 237 AA.
AC Q9NS33;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE HDCGC13P.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhao Z., Huang X., Li N., Zhu X., Cao X.;
RT "Novel gene identified from dendritic cells.";
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF067800; AAF75560.1; -.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
SQ SEQUENCE 237 AA; 27504 MW; 71197BFD337651EB CRC64;

Query Match 38.9%; Score 453; DB 4; Length 237;
Best Local Similarity 43.9%; Pred. No. 4e-38;
Matches 82; Conservative 34; Mismatches 67; Indels 4; Gaps 3;

QY 26 ISIALLSACFIVSCVVTYHFTYGETGKRLSELHSYHSLTCFSEGTKV--PAWGCCPASW 83
| : : ||||| : : : | : | : | : |
Db 52 IFELLALISFFIAFVIFQ-KYSQLEKKTTELVTLECVKKNMPVEETAWSCCPKNW 110
| : : ||||| : : : | : | : | : |

QY 84 KFGSSCYFISSEKVKSEQNCVEMGAHLVVFNTAEQNFIVQOLNESFSYFLGLSDP 143
| : : ||||| : : : | : | : | : |
Db 111 KFSNNCYFISTESASQWQSEKDCARMEAHLLVINTQEQDFIFQNLQEESEAYFVGLSDP 170
| : : ||||| : : : | : | : | : |

QY 144 QGNNNWQIDKTPYEKNVRFHGLGEPNHSACASIVFWK-PTGWNNDVICETRRNSIC 202
| : : ||||| : : : | : | : | : |
```



```
Best Local Similarity 68.1%; Pred. No. 7.6e-73;
Matches 145; Conservative 19; Mismatches 41; Indels 8; Gaps 4;

Qy 1 MMQEQQPQSTKRG--WLSRLMSVAGISIALLSACFIVSCVVTYHFTYGETGKRLSELH 58
Db 1 MVQERQSG---KGVCW-TLRLWSAAVISMLLSTCFIASCVVTYQFIMDQPSRRILYELH 56

Qy 59 SYHSSLTFCFSEGTGV--PANGCCPASWKSFGSGCYFISSEKWSKSEONCVEMGAHLVV 116
Db 57 TYHSSLTFCFSEGTGV--PANGCCPASWKSFGSGCYFISSEKWSKSEONCVEMGAHLVV 116

Qy 117 FNTAEQNFIVQOLNESFSYFLGLSDPQGNMNMOWIDKTPYEKNVRFWHLGEPNHSAEQC 176
Db 117 INTEAEQNFITQOLNESLSYFLGLSDPQGNMNMOWIDKTPYEKNVRFWHLGEPNHSAEQC 176

Qy 177 ASIVFWKPTGWNVDVICETRRNSICEMNKIYL 209
Db 177 VSIYVWNPWSKMGWNDVFCDSKHNSICEMKKIYL 209

RESULT 2
Q9JKF3 PRELIMINARY: PRT; 175 AA.
AC Q9JKF3
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Dectin-2 beta isoform.
GN CLCsf10.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C;
RA Arizumi K., Shen G., Shikano S., Ritter R. III, Zukas P.,
Edelbaum D., Morita A., Takashima A.;
RT "Cloning of a second dendritic cell-associated C-type lectin (dectin-2) and its alternatively spliced isoforms.";
RL J. Biol. Chem. 0:0-0(2000).
DR EMBL; AF240358; AAF67178.1; -.
DR HSSP; P22897; IEGG.
DR MGD; MGI:1861231; Clcfsf10.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; UNKNOWN_1.
DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
SQ SEQUENCE 175 AA; 20313 MW; EIC9G32687BF0EA9 CRC64;

Query Match 56.6%; Score 659; DB 11; Length 175;
Best Local Similarity 58.3%; Pred. No. 2.8e-59;
Matches 123; Conservative 16; Mismatches 34; Indels 38; Gaps 4;

Qy 1 MMQEQQPQSTKRG--WLSRLMSVAGISIALLSACFIVSCVVTYHFTYGETGKRLSELH 58
Db 1 MVQERQSG---KGVCW-TLRLWSAAVISMLLSTCFIASCV----- 38

Qy 59 SYHSSLTFCFSEGTGV--PANGCCPASWKSFGSGCYFISSEKWSKSEONCVEMGAHLVV 118
Db 39 -----EKMWGCCPNHWKSGSGCYLSTKENFWSTSEQNCVQMGAHLVVIN 84

Qy 119 TEAEQNFIVQOLNESFSYFLGLSDPQGNMNMOWIDKTPYEKNVRFWHLGEPNHSAEQAS 178
Db 85 TEAEQNFITQOLNESLSYFLGLSDPQGNMNMOWIDKTPYEKNVRFWHLGEPNHSAEQAS 178

Qy 179 IVFWKPTGWNVDVICETRRNSICEMNKIYL 209
Db 145 IVYVNPWSKMGWNDVFCDSKHNSICEMKKIYL 175

RESULT 3
Q9JKF2 PRELIMINARY: PRT; 168 AA.
AC Q9JKF2
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Dectin-2 gamma isoform.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

Q8WTT0 PRELIMINARY: PRT; 213 AA.
AC Q8WTT0
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Dendritic lectin (Blood dendritic cell antigen 2 protein).
GN CLCsf11 OR BDCA2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21426806; PubMed=11536172;
RA Arce I., Roda-Navarro P., Montoya M.C., Hernanz-Falcon P.,
Pulg-Kroger A., Fernandez-Ruiz E.;
RT "Molecular and genomic characterization of human DLEC, a novel member of the C-type lectin receptor gene family preferentially expressed on monocyte-derived dendritic cells.";
RT monocyte-derived dendritic cells.";
RL Eur. J. Immunol. 31:2733-2740(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Sohma Y., Johnston I., Dzionek A., Nagasaka T., Nagafune J., Okada T.,
Hirano T., Schmitz J., Yamaguchi Y.;
RT "Cloning and characterization of blood dendritic cell antigen 2 (BDCA-2), a C-type lectin expressed on plasmacytoid dendritic cells.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF325459; AAL37358.1; -.
DR EMBL; AF293615; AAL37036.1; -.
DR InterPro; IPR003654; GASA.
DR Pfam; PF02704; GASA; 1.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; UNKNOWN_1.
DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
KW Lectin.
SQ SEQUENCE 213 AA; 25037 MW; 5DC82C95BE2378C1 CRC64;

Query Match 54.7%; Score 637; DB 4; Length 213;
Best Local Similarity 54.0%; Pred. No. 6.2e-57;
Matches 115; Conservative 30; Mismatches 64; Indels 4; Gaps 2;

Qy 1 MMQEQQPQSTKRG--WLSRLMSVAGISIALLSACFIVSCVVTYHFTYGETGKRLSEL-- 57
Db 1 MVPEEPQDREKGLWFWLKVWMAVVSILLSVCTVSSVPHNPFMYSKTVKRLSKLRE 60

Qy 58 -HSYHSSLTFCFSEGTGV--PANGCCPASWKSFGSGCYFISSEKWSKSEONCVEMGAHLVV 116
Db 61 YQYHPSLTFCVMEGKDIEDWSCCPTPTSFQSSCYFISTGMOSWTQSKNCSVMGADLVV 120

Qy 117 FNTAEQNFIVQOLNESFSYFLGLSDPQGNMNMOWIDKTPYEKNVRFWHLGEPNHSAEQC 176
Db 121 INTREQDFIQLKNSSYFLGLSDPQGNMNMOWIDKTPYEKNVRFWHLGEPNHSAEQC 180

Qy 177 ASIVFWKPTGWNVDVICETRRNSICEMNKIYL 209
Db 181 ALINFRSSEWGWNDIHCHVPQKICKMKKIYI 213

RESULT 4
Q9JKF2 PRELIMINARY: PRT; 168 AA.
AC Q9JKF2
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Dectin-2 gamma isoform.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
```

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 22:55:56 ; Search time 30 seconds  
(without alignments)  
1435.462 Million cell updates/sec

Title: US-09-766-511B-53  
Perfect score: 1165  
Sequence: 1 MMQEQQPQSTKRGWLSIRL.....NDVICPRRNSICEMNKIYL 209

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21.\*

1: sp-archaea.\*

2: sp-bacteria.\*

3: sp-fungi.\*

4: sp-human.\*

5: sp-invertebrate.\*

6: sp-mammal.\*

7: sp-mhc.\*

8: sp-organelle.\*

9: sp-phage.\*

10: sp-plant.\*

11: sp-rodent.\*

12: sp-virus.\*

13: sp-vertebrate.\*

14: sp-unclassified.\*

15: sp-rvirus.\*

16: sp-bacteriap.\*

17: sp-archeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query %	Length	DB ID	Description
1	793	68.1	209	11 Q9JKF4	Q9jkf4 mus musculus
2	659	56.6	175	11 Q9JKF3	Q9jkf3 mus musculus
3	637	54.7	213	4 Q8WTT0	Q8wt0 homo sapien
4	592.5	50.9	168	11 Q9JKF2	Q9jkf2 mus musculus
5	534.5	45.9	182	4 Q8WXX8	Q8wx8 homo sapien
6	475.5	40.8	176	11 Q9D8Q7	Q9d8q7 mus musculus
7	458	39.3	237	4 Q9UT34	Q9uli34 homo sapien
8	458	39.3	237	4 Q9UMR7	Q9umr7 homo sapien
9	453	38.9	237	4 Q9NS33	Q9ns33 homo sapien
10	443	38.0	219	11 Q9Z2H6	Q9z2h6 mus musculus
11	441	37.9	215	4 Q8WX18	Q8wx18 homo sapien
12	436	37.4	238	11 Q9QZ15	Q9qz15 mus musculus
13	433.5	37.2	204	4 Q9H229	Q9h229 homo sapien
14	430	36.9	262	11 Q9Z3C7	Q9z3c7 mus musculus
15	425	36.5	230	4 Q8WXX9	Q8wxw9 homo sapien
16	398.5	34.2	219	4 Q9ULY5	Q9uly5 homo sapien

17	390.5	33.5	214	11 Q9E0Q8	Q9r0q8 mus musculus
18	254	21.8	236	13 Q8UUM9	Q8uum9 oryzias lat
19	250	21.5	292	4 Q14538	Q14538 homo sapien
20	249	21.4	742	4 Q9BYH7	Q9byh7 homo sapien
21	248	21.3	263	4 Q96QP3	Q96qp3 homo sapien
22	248	21.3	332	4 Q96QP5	Q96qp5 homo sapien
23	248	21.3	376	4 Q9BXS3	Q9bxs3 homo sapien
24	248	21.3	399	4 Q9H2X3	Q9h2x3 homo sapien
25	247	21.2	417	4 Q8TCR2	Q8tcr2 homo sapien
26	247	21.2	742	4 Q8WZA4	Q8wza4 homo sapien
27	246	21.1	237	11 Q91ZW8	Q91zw8 mus musculus
28	244	20.9	304	11 Q91YT3	Q91yt3 mus musculus
29	239	20.5	207	11 Q8VIK4	Q8vik4 mus musculus
30	238	20.4	399	4 Q9H2Q9	Q9hq9 homo sapien
31	236	20.3	312	4 Q96Q07	Q96q7 homo sapien
32	236	20.3	381	6 Q95LA8	Q95la8 macaca mula
33	236	20.3	381	6 Q8SQB2	Q8sqb2 macaca mula
34	235	20.2	268	4 Q96Q03	Q96q3 homo sapien
35	235	20.2	334	4 Q96QP9	Q96qp9 homo sapien
36	235	20.2	360	4 Q96Q04	Q96q4 homo sapien
37	235	20.2	380	4 Q96Q05	Q96q5 homo sapien
38	235	20.2	380	4 Q96Q00	Q96q0 homo sapien
39	235	20.2	381	6 Q95LC6	Q95lc6 macaca neme
40	235	20.2	404	4 Q96Q01	Q96q1 homo sapien
41	235	20.2	404	4 Q9NNX6	Q9nnx6 homo sapien
42	235	20.2	404	6 Q95JY6	Q95jy6 macaca mula
43	234.5	20.1	238	11 Q91ZX1	Q91zx1 mus musculus
44	234	20.1	404	6 Q95L98	Q95l98 pan troglod
45	229.5	19.7	208	11 Q91ZW7	Q91zw7 mus musculus

#### ALIGNMENTS

RESULT 1  
Q9JKF4  
ID Q9JKF4 PRELIMINARY; PRT; 209 AA.  
AC Q9JKF4;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Dectin-2 alpha isoform (C-type (calcium dependent, carbohydrate recognition domain) lectin, superfamily member 10).  
GN CLEGSF10.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BALB/C.  
RA Arizumi K., Shen G., Shikano S., Ritter R. III, Zukas P., Edelbaum D., Morita A., Takashima A.;  
RT "Cloning of a second dendritic cell-associated C-type lectin (dectin-2) and its alternatively spliced isoforms.";  
RL J. Biol. Chem. 0:0-0(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF240357; AAF67177.1; -.  
DR EMBL; BC023008; AAH23008.1; -.  
DR HSSP; P22897; 1EGG.  
DR MGD; MGI:1861231; Clecsf10.  
DR InterPro; IPR001304; Lectin\_C.  
DR Pfam; PF00059; lectin\_c; 1.  
DR SMART; SM00034; CLECT; 1.  
DR PROSITE; PS00615; C.TYPE.LECTIN\_1; UNKNOWN\_1.  
DR PROSITE; PS50041; C.TYPE.LECTIN\_2; 1.  
KW Lectin.  
SQ SEQUENCE 209 AA; 24324 MW; 7AB8BDC0DB1EA5ED CRC64;

Query Match 68.1%; Score 793; DB 11; Length 209;

Search completed: February 11, 2003, 23:01:57  
Job time : 30 secs

```
FT VARSPLIC 646 883 MISSING (IN GPI-ANCHORED ISOFORM).
FT CONFLICT 51 52 AL -> WV (IN REF. 3).
FT CONFLICT 503 503 V -> L (IN REF. 2).
FT CONFLICT 519 509 TV -> PA (IN REF. 2).
FT CONFLICT 526 526 G -> R (IN REF. 2).
FT CONFLICT 541 541 G -> A (IN REF. 2).
FT CONFLICT 556 556 R -> S (IN REF. 2).
FT CONFLICT 573 573 E -> A (IN REF. 2).
FT CONFLICT 583 583 V -> L (IN REF. 2).
FT CONFLICT 649 649 V -> L (IN REF. 2).
FT CONFLICT 670 670 P -> A (IN REF. 2).
FT CONFLICT 738 738 R -> A (IN REF. 2).
FT CONFLICT 809 809 R -> A (IN REF. 2).
SQ SEQUENCE 883 AA; 96057 MW; AC7ACC40CB53ED37 CRC64;

Query Match 15.7%; Score 183; DB 1; Length 883;
Best Local Similarity 26.1%; Pred. No. 7e-09;
Matches 62; Conservative 31; Mismatches 83; Indels 62; Gaps 13;

Qy 3 QEQQPQSTKRGWLSRLMSV-----AGISIALLSA-CFIVSCVVYHFTYG 48
Db 591 REVETPSEKSGRTVLGTGSVAQPLPTDSASRGVAVAPSGGDCIPSPC----- 631
Qy 49 ETGRKLSLHSHSLTCF--SEGTK---VPFWG---C-----CPASMKSPGSSCYFIS 94
Db 632 -----HNGGTCLEBEKEGFRCLCPVPGYGGDLCDVGLHPCSPGWEPFGACVKHF 679
Qy 95 SEEKVSKSEONCVEMGAHLVVFNTAEONFIVQOLNESFSYFLGLSDPGNNWQWIDK 154
Db 680 STRRSWEAESQCRALGALHTSICTPEQDFVNDYRE--YQWIGLNDRTIEGDFLWSDG 737
Qy 155 TP--YEKNVRVHILGEPNH-----SAEQACISVFWKPTGWNVDVICETRRNSICEMNKI 207
Db 738 PPLLYEN----WNPGQPDYSFLSGENCV-VMWIHDQG-QMSDVPNCVHLSYTKMGVLV 789

RESULT 15
ID PGCB_BOVIN STANDARD; PRT; 912 AA.
AC Q28062;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Brevican core protein precursor.
GN BCAN.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=94193597; PubMed=8144512;
RA Yamada H., Watanabe K., Shimonaka M., Yamaguchi Y.;
RT "Molecular cloning of brevicin, a novel brain proteoglycan of the
RT aggrecan/versican family.";
RL J. Biol. Chem. 269:10119-10126(1994).
CC -!- FUNCTION: MAY PLAY A ROLE IN THE TERMINALLY DIFFERENTIATING AND
CC THE ADULT NERVOUS SYSTEM DURING POSTNATAL DEVELOPMENT. COULD
CC STABILIZE INTERACTIONS BETWEEN HA AND BRAIN PROTEOGLYCAN.
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix (By
CC similarity).
CC -!- TISSUE SPECIFICITY: BRAIN; EXPRESSED IN CEREBELLAR ASTROCYTES BUT
CC NOT IN NEURONS.
CC -!- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE.
CC -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC -!- SIMILARITY: CONTAINS 2 LINK DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.
CC -!- SIMILARITY: BELONGS TO THE AGGREGAN/VERSICAN PROTEOGLYCAN FAMILY.
CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; X75887; CAA53481.1; -.
CC HSSP; P20893; LHLJ.
CC InterPro; IPR000561; EGF-like.
CC InterPro; IPR000742; EGF_2.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR003596; Ig_V.
CC InterPro; IPR001304; Lectin_C.
CC InterPro; IPR000538; Link.
CC InterPro; IPR000436; Sushi_SCR_CCP.
CC Pfam; PF00008; EGF; 1.
CC Pfam; PF00047; Ig; 1.
CC Pfam; PF00059; lectin_c; 1.
CC Pfam; PF00084; sushi; 1.
CC Pfam; PF00193; Xlink; 2.
CC ProDom; PD000918; Link; 2.
CC SMART; SM00032; CCP; 1.
CC SMART; SM00034; CLECT; 1.
CC SMART; SM00181; EGF; 1.
CC SMART; SM00406; IGV; 1.
CC SMART; SM00445; LINK; 2.
CC PROSITE; PS00022; EGF_1; 1.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS00290; Ig_MHC; 1.
CC PROSITE; PS01241; LINK; 2.
CC PROSITE; PS00615; C-TYPE LECTIN_1; 1.
CC PROSITE; PS50041; C-TYPE LECTIN_2; 1.
KW Glycoprotein; Hyaluronic acid; Proteoglycan; Lectin; Signal; Sushi;
KW EGF-like domain; Repeat; Immunoglobulin domain.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 912 BREVICAN CORE PROTEIN.
FT DOMAIN 32 158 IG-LIKE V-TYPE DOMAIN.
FT DOMAIN 174 251 LINK 1.
FT DOMAIN 272 353 LINK 2.
FT DOMAIN 647 683 EGF-LIKE.
FT DOMAIN 683 811 C-TYPE LECTIN.
FT DOMAIN 812 876 SUSHI.
FT DISULFID 57 137 BY SIMILARITY.
FT DISULFID 179 250 BY SIMILARITY.
FT DISULFID 203 224 BY SIMILARITY.
FT DISULFID 277 352 BY SIMILARITY.
FT DISULFID 301 322 BY SIMILARITY.
FT DISULFID 651 662 BY SIMILARITY.
FT DISULFID 656 671 BY SIMILARITY.
FT DISULFID 673 682 BY SIMILARITY.
FT DISULFID 689 700 BY SIMILARITY.
FT DISULFID 717 809 BY SIMILARITY.
FT DISULFID 785 801 BY SIMILARITY.
FT DISULFID 816 859 BY SIMILARITY.
FT DISULFID 845 872 BY SIMILARITY.
FT CARBOHYD 130 130 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 337 337 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 912 AA; 99554 MW; 677B3EB1C688C4D7 CRC64;

Query Match 15.7%; Score 182.5; DB 1; Length 912;
Best Local Similarity 29.5%; Pred. No. 8.1e-09;
Matches 49; Conservative 26; Mismatches 64; Indels 27; Gaps 10;

Qy 61 HSSLTCF--SEGTK---VPFWG---C-----CPASMKSPGSSCYFISSEKVKWSEQN 106
Db 657 HNGGTCLEEEGVRGCLCPGYGGDLCDVGLHFCSPGDWAFQACACYKHFARSWEARENK 716
Qy 107 CVEMGAHLVVFNTAEONFIVQOLNESFSYFLGLSDPGNNWQWIDKTP--YEKNVRWF 164
Db 717 CRMYGAHLASISTPEQDFINNRYRE--YQWIGLNDRTIEGDFLWSDGVPLLYEN---W 770
```

Query Match 15.7%; Score 183; DB 1; Length 883;  
Best Local Similarity 25.6%; Pred. No. 7e-09;  
Matches 61; Conservative 33; Mismatches 82; Indels 62; Gaps 13;

QY 3 QEQOPOSTKRGWLSRLWSV-----AGISIALLSA-CFIVSCVVVYHYTYG 48  
Db 581 RELLETSPEKSGRTVLAGTSVQAQVLPDTSASHGGVAVAPSSGDCIPSPC----- 631  
QY 49 ETGKRLSELHSYHSSITCF--SEGTK---VPAMG---C-----CPASWKSFGSSCYFIS 94  
Db 632 -----HNGTCLLEKEGFRCLCLPGYGGDLCDVGLHFCSPQWEAFQACAKYHP 679  
QY 95 SEKVMSKSEQNCQVEMGAHLVVFNTAEONFTVQOLNSESFYFGLSDPQGNNNQWIDK 154  
Db 680 STRSWEAEASQCRALGAHLTSICTPEQDFVNDYRE--YQWIGLNDRTIEGDFLWSDG 737  
QY 155 TP--YEKNVRFHLCENH---SACQASIVFWKFTGCHGNDVICETPRNSICEWNI 207  
Db 738 APLLYEN-----WNPQGPDSYFLSGENCV-VVMWHDQG-QWSDVPCNYHLSTCKMGLV 789

RESULT 14  
PCGB\_RAT  
ID PCGB\_RAT STANDARD; PRT: 883 AA.  
AC P55068; Q63040; Q62860; Q635113;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Brevican core protein precursor (Brain enriched hyaluronan binding protein) (BEHAV protein).  
GN BCAN.  
OC Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sprague-Dawley; TISSUE=Brain;  
RX MEDLINE=96070828; PubMed=7592978;  
RA Seidenbecher C.I., Richter K., Rauch U., Faessler R., Garner C.C., Gundelfinger E.D.;  
RT "Brevican, a chondroitin sulfate proteoglycan of rat brain, occurs as secreted and cell surface glycosylphosphatidylinositol-anchored isoforms.";  
RT J. Biol. Chem. 270:27206-27212(1995).  
RN [2]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 396-407.  
RC TISSUE=Brain;  
RX MEDLINE=96074575; PubMed=7488217;  
RA Yamada H., Watanabe K., Shimonaka M., Yamasaki M., Yamaguchi Y.;  
RT "cDNA cloning and the identification of an aggrecanase-like cleavage site in rat brevican.";  
RT Biochem. Biophys. Res. Commun. 216:957-963(1995).  
RN [3]  
RP SEQUENCE OF 1-423 FROM N.A.  
RC STRAIN=Sprague-Dawley; TISSUE=Brain;  
RX MEDLINE=94216386; PubMed=7512973;  
RA Jaworski D.M., Kelly G.W., Hockfield S.;  
RT "BEHAV, a new member of the proteoglycan tandem repeat family of hyaluronan-binding proteins that is restricted to the brain.";  
RT J. Cell Biol. 125:495-509(1994).  
RN [4]  
RP FUNCTION: MAY PLAY A ROLE IN THE TERMINALLY DIFFERENTIATING AND THE ADULT NERVOUS SYSTEM DURING POSTNATAL DEVELOPMENT. COULD STABILIZE INTERACTIONS BETWEEN HA AND BRAIN PROTEOGLYCAN. THE GPI-ANCHORED ISOFORM MAY FUNCTION AS A CHONDROITIN SULFATE-BEARING CELL SURFACE RECEPTOR.  
CC BEHAV is a member of the proteoglycan tandem repeat family of hyaluronan-binding proteins that is restricted to the brain.  
CC -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR MATRIX AND ONE FORM ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.  
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A SECRETED FORM (SHOWN HERE) AND A GPI-ANCHORED FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -!- TISSUE SPECIFICITY: BRAIN.  
CC -!- DEVELOPMENTAL STAGE: SOLUBLE FORM INCREASES FROM DAY P4 TO P64.  
CC

CC GPI-ANCHORED ISOFORM INCREASES AFTER DAY P8.  
CC -!- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE.  
CC -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.  
CC -!- SIMILARITY: CONTAINS 2 LINK DOMAINS.  
CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
CC -!- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.  
CC -!- SIMILARITY: BELONGS TO THE AGGREGAN/VERSICAN PROTEOGLYCAN FAMILY.  
CC -!- CAUTION: REF.3 SEQUENCE DIFFERS FROM THAT SHOWN FROM POSITION 364 ONWARD AND IS SMALLER (371 AA) DUE TO A FRAMESHIFT.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
CC  
CC EMBL: X79881; CAA56255.1; -;  
CC EMBL: X86406; CAA60160.1; -;  
CC EMBL: U37142; AAB87847.1; -;  
CC EMBL: Z28366; CAA82215.1; ALT\_FRAME.  
CC HSP: P20693; IHLJ.  
CC InterPro: IPR000361; EGF-like.  
CC InterPro: IPR000742; EGF\_2.  
CC InterPro: IPR003006; Ig\_MHC.  
CC InterPro: IPR003596; Ig\_v.  
CC InterPro: IPR001304; Lectin\_C.  
CC InterPro: IPR000538; Link.  
CC InterPro: IPR000436; Sushi\_SCR\_CCP.  
CC Pfam: PF00008; EGF; 1.  
CC Pfam: PF00047; Ig; 1.  
CC Pfam: PF00059; lectin\_c; 1.  
CC Pfam: PF00084; sushi; 1.  
CC Pfam: PF00193; Xlink; 2.  
CC ProDom: PD000918; Link; 2.  
CC SMART: SM00032; CCP; 1.  
CC SMART: SM00034; CLECT; 1.  
CC SMART: SM00181; EGF; 1.  
CC SMART: SM00406; IGV; 1.  
CC SMART: SM00445; LINK; 2.  
CC PROSITE: PS00022; EGF\_1; 1.  
CC PROSITE: PS01186; EGF\_2; 1.  
CC PROSITE: PS00290; Ig\_MHC; 1.  
CC PROSITE: PS01241; LINK; 2.  
CC PROSITE: PS00615; C\_TYPE\_LECTIN\_1; 1.  
CC PROSITE: PS50041; C\_TYPE\_LECTIN\_2; 1.  
CC KW Glycoprotein; Hyaluronic acid; Proteoglycan; Lectin; Signal; Sushi;  
CC KW EGF-like domain; Repeat; Immunoglobulin domain; Alternative splicing;  
CC  
CC SIGNAL 1 22 POTENTIAL.  
CC CHAIN 23 883 BREVICAN CORE PROTEIN.  
CC FT DOMAIN 32 157 IG-LIKE V-TYPE DOMAIN.  
CC FT DOMAIN 173 250 LINK 1.  
CC FT DOMAIN 271 352 LINK 2.  
CC FT DOMAIN 622 658 EGF-LIKE.  
CC FT DOMAIN 658 786 C-TYPE LECTIN.  
CC FT DOMAIN 787 851 SUSHI.  
CC FT DISULFID 56 136 BY SIMILARITY.  
CC FT DISULFID 178 249 BY SIMILARITY.  
CC FT DISULFID 202 223 BY SIMILARITY.  
CC FT DISULFID 276 351 BY SIMILARITY.  
CC FT DISULFID 300 321 BY SIMILARITY.  
CC FT DISULFID 626 637 BY SIMILARITY.  
CC FT DISULFID 648 657 BY SIMILARITY.  
CC FT DISULFID 791 834 BY SIMILARITY.  
CC FT DISULFID 820 847 BY SIMILARITY.  
CC FT CARBOHYD 129 129 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC FT VARSPLIC 625 645 DCIPSPNGGTCLKEEGFR -> NSAEGSMPAFLFLLL OLWDT (IN GPI-ANCHORED ISOFORM).  
CC

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

```

CC EMBL: D88577; BAA13647.1; -.
CC HSP: P20693; 1HLJ.
CC MGD: MGI:1859834; Clecsf13.
CC InterPro: IPR001304; Lectin_C.
CC Pfam: PF00059; lectin_c; 1.
CC SMART: SM00034; CLECT; 1.
CC PROSITE: PS00615; C-TYPE_LECTIN_1; 1.
CC PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
KW Receptor; Transmembrane; Glycoprotein; Lectin; Signal-anchor;
KW Endocytosis.
FT DOMAIN 1 42 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 43 69 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
(POTENTIAL).
FT DOMAIN 70 548 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 438 538 C-TYPE LECTIN (SHORT FORM).
FT DISULFID 440 536 BY SIMILARITY.
FT DISULFID 516 528 BY SIMILARITY.
FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 92 92 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 115 115 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 132 132 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 209 209 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 548 AA; 61268 MW; 6F6495E820E73BD9 CRC64;

```

Query Match 16.3%; Score 189.5; DB 1; Length 548;

Best Local Similarity 29.2%; Pred. No. 1.1e-09;

Matches 38; Conservative 26; Mismatches 51; Indels 15; Gaps 4;

```

QY 82 SWKSGSCYFISSEKRWKSEQNCVEMGAHLWVFNTAEQNFIVQQLNESFSYFLGLS 141
   ||| | : : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 415 NWKYFNGFYFSDKPKWAEKFKCTSGAHLASVTSQEAFLV-QITSSGDHWIGLT 473

QY 142 DPQGNWQWIDKTPYE--KNVRFHLGEP-----NHSAEQCAISVFWKPTGWGNDVI 193
   | : | : | : : : | : : | : : | : : | : : | : : | : : | : : |
Db 474 DQGTGLRWVWDGTPFNNAQSGKFWGKNQPDNWRHNRGEREDCVHV-----RQWNWMA 527

QY 194 CETRNSICE 203

Db 528 CGSSYPWVCK 537

```

```

RESULT 13
ID_PGCB_MOUSE STANDARD; PRT; 883 AA.
AC Q61361;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Brevican core protein precursor.
GN BCAN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Brain;
RX MEDLINE=97432816; PubMed=9286696;
RA Rauch U., Meyer H., Brakebusch C., Seidenbecher C., Gundelfinger E.D.,
RA Beier D.R., Fassler R.;
RT "Sequence and chromosomal localization of the mouse brevican gene.";
RL Genomics 44:15-21(1997).
CC -!- FUNCTION: MAY PLAY A ROLE IN THE TERMINALLY DIFFERENTIATING AND
      THE ADULT NERVOUS SYSTEM DURING POSTNATAL DEVELOPMENT. COULD

```

```

CC STABILIZE INTERACTIONS BETWEEN HA AND BRAIN PROTEOGLYCAN.
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix (By
CC similarity).
CC -!- TISSUE SPECIFICITY: BRAIN (BY SIMILARITY).
CC -!- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC -!- SIMILARITY: CONTAINS 2 LINK DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.
CC -!- SIMILARITY: BELONGS TO THE AGGREGAN/VERSICAN PROTEOGLYCAN FAMILY.

```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

```

CC EMBL: X87096; CAA60575.1; -.
CC HSP: P20693; 1HLJ.
CC MGD: MGI:1096385; Bcan.
CC InterPro: IPR000561; EGF-like.
CC InterPro: IPR000742; EGF_2.
CC InterPro: IPR003006; Ig_MHC.
CC InterPro: IPR003596; Ig_v.
CC InterPro: IPR001304; Lectin_C.
CC InterPro: IPR000538; Link.
CC InterPro: IPR000436; Sushi_SCR_CCP.
CC Pfam: PF00008; EGF; 1.
CC Pfam: PF00047; Ig; 1.
CC Pfam: PF00059; lectin_c; 1.
CC Pfam: PF00084; sushi; 1.
CC Pfam: PF00193; Xlink; 2.
CC ProDom: PD000918; Link; 2.
CC SMART: SM00032; CCP; 1.
CC SMART: SM00034; CLECT; 1.
CC SMART: SM00181; EGF; 1.
CC SMART: SM00406; IGV; 1.
CC SMART: SM00445; Link; 2.
CC PROSITE: PS00022; EGF_1; 1.
CC PROSITE: PS01186; EGF_2; 1.
CC PROSITE: PS00290; IG_MHC; 1.
CC PROSITE: PS01241; LINK; 2.
CC PROSITE: PS00615; C-TYPE_LECTIN_1; 1.
CC PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
KW Glycoprotein; Hyaluronic acid; Proteoglycan; Lectin; Signal; Sushi;
KW EGF-like domain; Repeat; Immunoglobulin domain.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 883 BREVICAN CORE PROTEIN.
FT DOMAIN 32 157 IG-LIKE V-TYPE DOMAIN.
FT DOMAIN 173 250 LINK 1.
FT DOMAIN 271 352 LINK 2.
FT DOMAIN 622 658 EGF-LIKE.
FT DOMAIN 658 786 C-TYPE LECTIN.
FT DOMAIN 787 851 SUSHI.
FT DISULFID 56 136 BY SIMILARITY.
FT DISULFID 178 249 BY SIMILARITY.
FT DISULFID 202 223 BY SIMILARITY.
FT DISULFID 276 351 BY SIMILARITY.
FT DISULFID 300 321 BY SIMILARITY.
FT DISULFID 626 637 BY SIMILARITY.
FT DISULFID 631 646 BY SIMILARITY.
FT DISULFID 648 657 BY SIMILARITY.
FT DISULFID 664 675 BY SIMILARITY.
FT DISULFID 692 784 BY SIMILARITY.
FT DISULFID 760 776 BY SIMILARITY.
FT DISULFID 791 834 BY SIMILARITY.
FT DISULFID 820 847 BY SIMILARITY.
FT CARBOHYD 129 129 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 883 AA; 96013 MW; CC2C3C97B453E45 CRC64;

```



CC	Anthochordaris.
OX	NCBI_TaxID=7629;
RN	[1]
RP	SEQUENCE.
RX	MEDLINE=87194838; PubMed=3571253;
RA	Giga Y., Ikal A., Takahashi K.;
RT	"The complete amino acid sequence of echinoidin, a lectin from the
RT	coelomic fluid of the sea urchin Anthochordaris crassispina. Homologies
RT	with mammalian and insect lectins.";
RL	J. Biol. Chem. 262:6197-6203(1987).
CC	-!- FUNCTION: ROLE IN THE DEFENCE SYSTEM OF THE ORGANISM AGAINST
CC	MICROORGANISMS. THIS LECTIN IS SPECIFIC TO GAL-GAL(NAC).
CC	-!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
CC	-!- TISSUE SPECIFICITY: COELOMIC FLUID.
CC	-!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
DR	PIR; A26697; A26697.
DR	HSP; P20693; 1HLJ.
DR	InterPro: IPR001304; Lectin_c.
DR	Pfam: PF000059; lectin_c; 1.
DR	SMART; SM000334; CLECT; 1.
DR	PROSITE; PS00615; C-TYPE_LECTIN_1; 1.
DR	PROSITE; PS00041; C-TYPE_LECTIN_2; 1.
KW	Glycoprotein; Lectin.
FT	DOMAIN 1 143 C-TYPE LECTIN (LONG FORM).
FT	DISULFID 2 2 INTERCHAIN.
FT	DISULFID 3 14
FT	DISULFID 31 141
FT	DISULFID 116 132
FT	CARBOHYD 38 38
FT	SITE 39 41 O-LINKED.
FT	CELL ATTACHMENT SITE (POTENTIAL).
SQ	SEQUENCE 147 AA; 16661 MW; 849E98645481A062 CRC64;
Query Match 16.7%; Score 194.5; DB 1; Length 147;	
Best Local Similarity 31.3%; Pred. No. 8.3e-11;	
Matches 46; Conservative 20; Mismatches 58; Indels 23; Gaps	
Qy	77 GCCPASPASKGSCGYCFYSSEKVVSKSEQNC-----VEMGAHLVVFNTEAQNFI 1256
Db	
Db	1 GCCPFTFTSGSCNYCFRFAVSLTWAEQECQSFSPSGRDIDISIGHLSIHSETQNFV 60
Qy	127 VQQLNESF-----SYFLGLSDPQNNWQIDKTPYEKNVRFWHLGEPNH--SAEQCAS 178
Db	
Db	61 YHYFETSTKDDTTPMMLGFNDRTTEGFWQTDGSP--NDFTAWGSPNDNYSGEDCTQ 118
Qy	179 IVFWKPTGWNVDVIC-ETRRNSTCEM 204
Db	
Db	119 MVM--GAGLAWIDLPSCSRYHLICKL 143
RESULT 12	
KUCR_MOUSE	
ID	KUCR_MOUSE STANDARD; PRT; 548 AA.
AC	P70194;
DT	01-NOV-1997 (Rel. 35, Created)
DT	01-NOV-1997 (Rel. 35, Last sequence update)
DT	15-JUN-2002 (Rel. 41, Last annotation update)
DE	C-type lectin 13 (Kupffer cell receptor).
GN	CLEC3F13 OR KCLR.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	STRAIN=BALB/c; TISSUE=Liver;
RA	Takezawa R., Magatsuma H., Nomoto C., Watanabe Y., Akaike T.;
RL	Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC	-!- FUNCTION: RECEPTOR WITH AN AFFINITY FOR GALACTOSE AND FUCOSE.
CC	COULD BE INVOLVED IN ENOCYTOSIS.
CC	-!- SUBCELLULAR LOCATION: Type II membrane protein.
CC	-!- TISSUE SPECIFICITY: KUPFFER CELLS.
CC	-!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN

Db 260 KQPDNWNHGLGGEDCAHFT----SDGRWDDVCQRPYRWCEM 300

# RESULT 9

```

ID  LECH_MOUSE          STANDARD;      PRT;    283 AA.
AC  P34927; O64363;
DT  01-FEB-1994 (Rel. 28, Created)
DT  15-JUL-1998 (Rel. 36, Last sequence update)
DT  30-MAY-2000 (Rel. 39, Last annotation update)
DE  Asialoglycoprotein receptor 1 (Hepatic lectin 1) (MHL-1) (ASGP-R)
DE  (ASGPR).
DE  ASGR1 OR ASGR-1.
OS  Mus musculus (Mouse).
OC  Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=93176818; PubMed=8439566;
RA  Takezawa R., Shirazawa K., Watanabe Y., Akaiki T.;
RT  "Determination of mouse major asialoglycoprotein receptor cDNA
RT  sequence.";
RL  Biochim. Biophys. Acta 1172:220-222(1993).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=BALB/C; TISSUE=Liver;
RX  MEDLINE=95047431; PubMed=7958950;
RA  Monroe R.S., Huber B.E.;
RT  "The major form of the murine asialoglycoprotein receptor: cDNA
RT  sequence and expression in liver, testis and epididymis.";
RL  Gene 148:237-244(1994).
CC  -!- FUNCTION: MEDIATES THE ENDOCYTOSIS OF PLASMA GLYCOPROTEINS TO
CC  WHICH THE TERMINAL SIALIC ACID RESIDUE ON THEIR COMPLEX
CC  CARBOHYDRATE MOETIES HAS BEEN REMOVED. THE RECEPTOR RECOGNIZES
CC  TERMINAL GALACTOSE AND N-ACETYL GALACTOSAMINE UNITS. AFTER LIGAND
CC  BINDING TO THE RECEPTOR, THE RESULTING COMPLEX IS INTERNALIZED AND
CC  TRANSPORTED TO A SORTING ORGANELLE, WHERE RECEPTOR AND LIGAND ARE
CC  DISASSOCIATED. THE RECEPTOR THEN RETURNS TO THE CELL MEMBRANE
CC  SURFACE.
CC  -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC  -!- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN HEPATIC PARENCHYMAL
CC  CELLS.
CC  -!- MISCELLANEOUS: CALCIUM IS REQUIRED FOR LIGAND BINDING.
CC  -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; D13517; BAA02734.1; -
DR  DR  EMBL; U09362; AAB60441.1; -
DR  DR  EMBL; U08372; AAB60440.1; -
DR  PIR; S29855; S29855.
DR  HSP; P20693; 1HLJ.
DR  MGD; MGI:88081; Asgr1.
DR  InterPro; IPR001304; Lectin_C.
DR  Pfam; PF00059; lectin_c; 1.
DR  SMART; SM00034; CLECT; 1.
DR  PROSITE; PS00615; C-TYPE LECTIN_1; 1.
DR  PROSITE; PS50041; C-TYPE LECTIN_2; 1.
KW  Lectin; Glycoprotein; Receptor; Endocytosis; Transmembrane;
KW  Calcium; Signal-anchor; Phosphorylation.
FT  INIT MET 0 0 BY SIMILARITY.
FT  DOMAIN 1 38 CYTOPLASMIC (POTENTIAL).
FT  TRANSMEM 39 59 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT  (POTENTIAL).
FT  DOMAIN 60 283 EXTRACELLULAR (POTENTIAL).
FT  DOMAIN 151 277 C-TYPE LECTIN (LONG FORM).

```

```

FT SITE 4 7 ENDOCYTOSIS SIGNAL (POTENTIAL).
FT DISULFID 152 163 BY SIMILARITY.
FT DISULFID 180 275 BY SIMILARITY.
FT DISULFID 253 267 BY SIMILARITY.
FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 77 77 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 145 145 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 150 150 I -> T (IN REF. 1).
SQ SEQUENCE 283 AA; 32472 MW; 982A5D305AAE0D8F CRC64;

Query Match 18.6%; Score 217; DB 1; Length 283;
Best Local Similarity 31.6%; Pred. No. 1.6e-12;
Matches 42; Conservative 25; Mismatches 52; Indels 14; Gaps 4;

```

```

QY 78 CCPSAKMSKSSCYFISSEKVKSKSFQNCVEMGAHLVVFENTEAEQNFVQQLNESFSYF 137
Db 151 CCPINWVEGSCYWFSSSRPMTADKYCOLENAHLVVVTSRDEQNFQRMHG-PLNTW 209
QY 138 LGLSDPOGNNNNWOWIDKTPYKKNVRFWHLGEPNH-----SAEQCAISIVFWKPTGCGWN 190
Db 210 IGLTDQ--NGPKWVDGTDYETGFQNRWPPQPDNWNHGLGGEDCAHFT----TDGRWN 263
QY 191 DVCETFRNSICE 203
Db 264 DDVCRRPYRWCE 276

```

```

RESULT 10
LECH_HUMAN          STANDARD;      PRT;    290 AA.
ID  LECH_HUMAN
DT  01-APR-1988 (Rel. 07, Created)
DT  01-APR-1988 (Rel. 07, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  Asialoglycoprotein receptor 1 (Hepatic lectin H1) (ASGPR) (ASGP-R).
GN  ASGR1.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=85130911; PubMed=2982798;
RA  Spiess M., Schwartz A.L., Lodish H.F.;
RT  "Sequence of human asialoglycoprotein receptor cDNA. An internal
RT  signal sequence for membrane insertion.";
RL  J. Biol. Chem. 260:1979-1982(1985).
RN  [2]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=86079574; PubMed=3753585;
RA  Spiess M., Lodish F.;
RT  "An internal signal sequence: the asialoglycoprotein receptor membrane
RT  anchor.";
RL  Cell 44:177-185(1986).
RN  [3]
RP  SEQUENCE FROM N.A.
RA  Wang H., Gao X., Li L., Lou H., Huang Y., Wang B., Han J.;
RT  "Human asialoglycoprotein receptor 1 gene is expressed in SH-SY5Y
RT  human neuroblastoma cells.";
RL  Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC  -!- FUNCTION: MEDIATES THE ENDOCYTOSIS OF PLASMA GLYCOPROTEINS TO
CC  WHICH THE TERMINAL SIALIC ACID RESIDUE ON THEIR COMPLEX
CC  CARBOHYDRATE MOETIES HAS BEEN REMOVED. THE RECEPTOR RECOGNIZES
CC  TERMINAL GALACTOSE AND N-ACETYL GALACTOSAMINE UNITS. AFTER LIGAND
CC  BINDING TO THE RECEPTOR, THE RESULTING COMPLEX IS INTERNALIZED AND
CC  TRANSPORTED TO A SORTING ORGANELLE, WHERE RECEPTOR AND LIGAND ARE
CC  DISASSOCIATED. THE RECEPTOR THEN RETURNS TO THE CELL MEMBRANE
CC  SURFACE.
CC  -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC  -!- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN HEPATIC PARENCHYMAL
CC  CELLS.
CC  -!- PTM: PHOSPHORYLATED ON A CYTOPLASMIC SER RESIDUE.
CC  -!- MISCELLANEOUS: CALCIUM IS REQUIRED FOR LIGAND BINDING.

```

```

CC CC SURFACE.
CC CC -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC CC -!- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN HEPATIC PARENCHYMAL
CC CC CELLS.
CC CC -!- MISCELLANEOUS: CALCIUM IS REQUIRED FOR LIGAND BINDING.
CC CC -!- MISCELLANEOUS: TWO TYPES OF RAT HEPATIC LECTIN HAVE BEEN
CC CC IDENTIFIED, RHL-1 AND RHL-2/3, HAVING A RELATIVE ABUNDANCE OF 4:1.
CC CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL: K02817; AAA42037.1; -.
CC CC EMBL: M21770; AAA40764.1; -.
CC CC PIR: A03166; LNRTL.
CC CC HSSP: P20693; 1HLJ.
CC CC InterPro: IPR001304; Lectin_C.
CC CC SMART: SM00034; lectin_c; 1.
CC CC PROSITE: PS00615; C-TYPE-LECTIN_1; 1.
CC CC PROSITE: PS50041; C-TYPE-LECTIN_2; 1.
CC CC LECTIN; Glycoprotein; Receptor; Endocytosis; Transmembrane;
CC CC Calcium; Signal-anchor; Phosphorylation.
CC CC INIT MET 0 0
CC CC DOMAIN 1 38
CC CC TRANSMEM 39 59
CC CC CYTOPLASMIC (POTENTIAL).
CC CC SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
CC CC (POTENTIAL).
CC CC EXTRACELLULAR (POTENTIAL).
CC CC C-TYPE LECTIN (LONG FORM).
CC CC ENDOCYTOSIS SIGNAL (POTENTIAL).
CC CC BY SIMILARITY.
CC CC BY SIMILARITY.
CC CC BY SIMILARITY.
CC CC BY SIMILARITY.
CC CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC O -> R (IN REF. 2).
CC CC CONFLICT 60 60
CC CC SEQUENCE 283 AA; 32718 MW; 3BA2631A5E28A993 CRC64;
CC CC -----
Query Match 19.4%; Score 226.5; DB 1; Length 283;
Best Local Similarity 25.9%; Pred. No. 2.3e-13;
Matches 64; Conservative 31; Mismatches 69; Indels 83; Gaps 11;
QY 25 GISIALLSACFVSCVTVTH-----FTYGE-TGKRL 54
DB 45 GLSILL-----VVVCVITSONSQLRDLRLVRQNFSTVSTEDQVKALTTQGERVGRKM 100
QY 55 -----SELHSYHS-----SLTCFSEGGKVPWG-----CCPASW 83
DB 101 KLVESOLEKHQEDLRHDSRLLLHVKQLVSDVRSUSQMAALR-----GNGSERICCPINW 156
QY 84 KSFSGSYFISBEKVKWSEQNCVEMGAHLVVFNTAEQNFIVQOLNESFSYFLGLSDP 143
DB 157 VEYEGSGYFWSSSVKPTWTEADKYQLENAHLVVVTSWEQRFVQOHMG-PLNTWIGLTDQ 215
QY 144 QGNKNWQIDKTPYEKNVRFWHLGEPNH-----SAEQACISVFWKPTGWNVDVICET 196
DB 216 --NGPKWVDGTDYETGFKNWRPFGDPDDWYGHGLGGEDCAHFT----TDGHNDDVCR 269
QY 197 RRNSICE 203
DB 270 PYRWVCE 276

```

RESULT 8  
MMGL\_RAT  
ID MMGL\_RAT PRT; 306 AA.  
AC P49301;

```

DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Macrophage asialoglycoprotein-binding protein (M-ASGP-BP) (Macrophage
DE galactose/N-acetylgalactosamine-specific lectin) (MMGL).
GN MGL.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=90293078; PubMed=2358462;
RA Ii M., Kurata H., Itoh N., Yamashina I., Kawasaki T.;
RT "Molecular cloning and sequence analysis of cDNA encoding the
RT macrophage lectin specific for galactose and N-acetylgalactosamine.";
RL J. Biol. Chem. 265:11295-11298(1990).
RN [2]
RP PRELIMINARY SEQUENCE OF 9-28.
RX MEDLINE=88339956; PubMed=3421964;
RA Ii M., Kawasaki T., Yamashina I.;
RT "Structural similarity between the macrophage lectin specific for
RT galactose/N-acetylgalactosamine and the hepatic asialoglycoprotein
RT binding protein.";
RL Biochem. Biophys. Res. Commun. 155:720-725(1988).
CC -!- FUNCTION: RECOGNIZES TERMINAL GALACTOSE AND N-ACETYL GALACTOSAMINE
CC UNITS.
CC -!- SUBUNIT: MONOMER.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL: J05495; AAA41216.1; -.
CC CC HSSP: P06734; 1HLI.
CC CC InterPro: IPR001304; Lectin_C.
CC CC Pfam: PF000059; lectin_c; 1.
CC CC SMART: SM00034; CLECT; 1.
CC CC PROSITE: PS00615; C-TYPE-LECTIN_1; 1.
CC CC PROSITE: PS50041; C-TYPE-LECTIN_2; 1.
CC CC LECTIN; Glycoprotein; Transmembrane; Calcium; Signal-anchor.
CC CC DOMAIN 1 37
CC CC TRANSMEM 38 58
CC CC CYTOPLASMIC (POTENTIAL).
CC CC SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
CC CC (POTENTIAL).
CC CC EXTRACELLULAR (POTENTIAL).
CC CC C-TYPE LECTIN (LONG FORM).
CC CC BY SIMILARITY.
CC CC BY SIMILARITY.
CC CC BY SIMILARITY.
CC CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CC SEQUENCE 306 AA; 34242 MW; D68A5DFF0B9E8F13 CRC64;
CC CC -----
Query Match 19.1%; Score 222; DB 1; Length 306;
Best Local Similarity 31.5%; Pred. No. 6.4e-13;
Matches 52; Conservative 25; Mismatches 72; Indels 16; Gaps 5;
QY 49 BTGKRLSELHSYHSLSLTCFSEGGK-VPWAGCCPASWKFSGSSCYFISSEKVKWSEQN 106
DB 143 EITDRVQOLGKDLKTLTLCQLASLKNNGSAVACCPHLHMEHGGSCYWFSSQCKPWPEADKY 202
QY 107 CVMGAHLVVFNTAEQNFIVQOLNESFSYFLGLSDPQGNNNWQIDKTPYEKNVRFWHL 166
DB 203 QOLENSNLVVVNSLAEQNFILQTHMG-SVVTWIGLTDQ--NGPWRWVDGTDYKGFTHWAP 259
QY 167 GEPNH-----SAEQACISVFWKPTGWNVDVICETRRNSICEM 204

```

Db 505 NWRHNGEREDCVHLQRM-----WMDMACGTAYNWVCK 537

## RESULT 6

## LECH\_CHICK

ID LECH\_CHICK STANDARD; PRT; 207 AA.

AC P02707;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE Hepatic lectin.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.

OX NCBI\_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=91268022; PubMed=2050668;

RA Bezouska K., Crichlow G.V., Rose J.M., Taylor M.E., Drickamer K.;

RT "Evolutionary conservation of intron position in a subfamily of genes

RT encoding carbohydrate-recognition domains.";

RL J. Biol. Chem. 266:11604-11609(1991).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=88186849; PubMed=3281941;

RA Mellow T.E., Halberg D., Drickamer K.;

RT "Endocytosis of N-acetylglucosamine-containing glycoproteins by rat

RT fibroblasts expressing a single species of chicken liver glycoprotein

RT receptor.";

RL J. Biol. Chem. 263:5468-5473(1988).

RN [3]

RP SEQUENCE

RX MEDLINE=81215504; PubMed=7240175;

RA Drickamer K.;

RT "Complete amino acid sequence of a membrane receptor for

RT glycoproteins. Sequence of the chicken hepatic lectin.";

RL J. Biol. Chem. 256:5827-5839(1981).

CC -!- FUNCTION: HEPATIC LECTIN IS A MEMBRANE RECEPTOR PROTEIN THAT

CC RECOGNIZES AND BINDS EXPOSED N-ACETYLGLUCOSAMINE MOIETIES OF

CC PLASMA GLYCOPROTEINS, THUS MEDIATING THEIR CLEARANCE (FROM THE

CC CIRCULATION) AND ENDOCYTOSIS.

CC -!- SUBCELLULAR LOCATION: Type II membrane protein.

CC -!- PTM: SOME OR ALL OF THE CYSTEINES ARE INVOLVED IN DISULFIDE BONDS.

CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; M63230; AAA48814.1; .

DR EMBL; M63225; AAA48814.1; JOINED.

DR EMBL; M63226; AAA48814.1; JOINED.

DR EMBL; M63227; AAA48814.1; JOINED.

DR EMBL; M63228; AAA48814.1; JOINED.

DR EMBL; M63229; AAA48814.1; JOINED.

DR EMBL; J03188; AAA48937.1; .

DR PIR; A03167; LNCHL.

DR HSSP; P20693; LHLJ.

DR InterPro; IPR001304; Lectin\_C.

DR Pfam; PF00059; lectin\_c; 1.

DR SMART; SM00034; CLECT; 1.

DR PROSITE; PS00615; C-TYPE LECTIN.1; 1.

DR PROSITE; PS00041; C-TYPE LECTIN.2; 1.

KW Lectin; Glycoprotein; Receptor; Endocytosis; Transmembrane;

KW Signal-anchor; Acetylation.

FT MOD\_RES 1 1

FT DOMAIN 1 23

FT CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 24 48 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
FT FT EXTRACELLULAR (POTENTIAL).  
FT DOMAIN 49 207 C-TYPE LECTIN (LONG FORM).  
FT DOMAIN 77 203 BY SIMILARITY.  
FT DISULFID 78 92 BY SIMILARITY.  
FT DISULFID 109 201 BY SIMILARITY.  
FT DISULFID 179 193 BY SIMILARITY.  
FT CARBOHYD 67 67 N-LINKED (GLCNAC...).  
SQ SEQUENCE 207 AA; 24216 MW; 1F6B36FDB32899DB CRC64;

Query Match 19.8%; Score 231; DB 1; Length 207;

Best Local Similarity 29.6%; Pred. No. 6.3e-14;

Matches 56; Conservative 29; Mismatches 78; Indels 26; Gaps 7;

QY 18 LRLWSVAGISIALLSACFIYSCVYHYFTYGTGKRLSELHSYHSSITCFSEGTKYPAWG 77

Db 37 LLTLTSSVSLARIA-----LSSKLSTLQSEPKHNFSSRDSLLFPFC-G 79

QY 78 CCPASWKSFGSSCYFISSEKQVMSKQNCVEMGAHLVVFTEAEQNFIV-QOLNESFSY 136

Db 80 AQSQWEYFEGRCYCYFSLRSMWHKAKAEEMHSHLIIDSYAKQNFVYFTRNERF-- 137

QY 137 FLGLSDPQGNMOWIDKTPYKXNVRFWHLGEPNHS--EQCASIVFWKPTGWNWDVIC 194

Db 138 WIGLTDENQEGEOWVDGTDTRSFTFWKGEPPNRCFNEDCAHV--W--TSGOWNDVYC 193

QY 195 ETRNSICE 203

Db 194 TYECYVYCE 202

## RESULT 7

## LECH\_RAT

ID LECH\_RAT STANDARD; PRT; 283 AA.

AC P02706;

DT 21-JUL-1986 (Rel. 01, Created)

DT 13-AUG-1987 (Rel. 05, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE Asialoglycoprotein receptor 1 (Hepatic lectin 1) (RHL-1) (ASGP-R)

DE (ASGPR).

GN ASGR1 OR ASGR-1.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=86008335; PubMed=2995379;

RA Leung J.O., Holland E.C., Drickamer K.;

RT "Characterization of the gene encoding the major rat liver

RT asialoglycoprotein receptor.";

RL J. Biol. Chem. 260:12523-12527(1985).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=85063786; PubMed=6095287;

RA Holland E.C., Leung J.O., Drickamer K.;

RT "Rat liver asialoglycoprotein receptor lacks a cleavable NH2-terminal

RT signal sequence.";

RL Proc. Natl. Acad. Sci. U.S.A. 81:7338-7342(1984).

RN [3]

RP SEQUENCE OF 11-283 FROM N.A.

RX MEDLINE=87026895; PubMed=2945599;

RA Watts C.;

RT "Isolation and expression of cDNA clones for a rat liver

RT asialoglycoprotein receptor.";

RL Biosci. Rep. 6:527-534(1986).

CC -!- FUNCTION: MEDIATES THE ENDOCYTOSIS OF PLASMA GLYCOPROTEINS TO

CC WHICH THE TERMINAL SIALIC ACID RESIDUE ON THEIR COMPLEX

CC CARBOHYDRATE MOIETIES HAS BEEN REMOVED. THE RECEPTOR RECOGNIZES

CC TERMINAL GALACTOSE AND N-ACETYLGLUCOSAMINE UNITS. AFTER LIGAND

CC BINDING TO THE RECEPTOR, THE RESULTING COMPLEX IS INTERNALIZED AND

CC TRANSPORTED TO A SORTING ORGANELLE, WHERE RECEPTOR AND LIGAND ARE

CC DISASSOCIATED. THE RECEPTOR THEN RETURNS TO THE CELL MEMBRANE

```

[2]
RN SEQUENCE OF 102-120 AND 137-151.
RC STRAIN-C3H/HeN;
RX MEDLINE=89197865; PubMed=3241002;
RA Oda S., Sato M., Toyoshima S., Osawa T.;
RT "Purification and characterization of a lectin-like molecule specific
RT for galactose/N-acetyl-galactosamine from tumoricidal macrophages.";
RL J. Biochem. 104:600-605(1988).
CC -!- FUNCTION: RECOGNIZES TERMINAL GALACTOSE AND N-ACETYLGALACTOSAMINE
CC UNITS MAY PARTICIPATE IN THE INTERACTION BETWEEN TUMORICIDAL
CC MACROPHAGES AND TUMOR CELLS.
CC
CC -!- SUBUNIT: HOMO-OLIGOMER.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC -!- TISSUE SPECIFICITY: IS EXPRESSED ON THE SURFACE OF ACTIVATED
CC MACROPHAGES.
CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S36676; AAB22171.1; -.
DR HSSP; P06734; 1HLI.
DR MGD; MGI:96975; Mgi.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
DR Lectin; Glycoprotein; Transmembrane; Calcium; Signal-anchor.
KW DOMAIN 1 35 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 36 56 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT FT
FT DOMAIN 57 304
FT FT EXTRACELLULAR (POTENTIAL).
FT DOMAIN 172 298 C-TYPE LECTIN (LONG FORM).
FT DISULFID 173 184 BY SIMILARITY.
FT DISULFID 201 296 BY SIMILARITY.
FT DISULFID 274 288 BY SIMILARITY.
FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 304 AA; 34596 MW; 3F79CD12C34F5BCC CRC64;

Query Match 20.9%; Score 244; DB 1; Length 304;
Best Local Similarity 33.7%; Pred. No. 6.6e-15;
Matches 55; Conservative 27; Mismatches 59; Indels 22; Gaps 6;

QY 54 LSELHSHSLTCTCFSEGTGVPAWGCCPASWKSFGSSCYFISSEKVKWSKSEQNCV 108
DB 146 VOOLRRKDKALTCQLANLKNNGSEV---ACPLHWTEHSGSCYWFSEKSWPEADKYCR 202
QY 109 EMGAHLVFNTEAEQNTFVQQLNESFSYFLGLSDPQGNNNWQIDKTP--YKKNVRFHILGE 168
DB 203 LENSLLVNVNSLEEQNFQNLNANVSN-IGLTQD--NGPWRVVDGDTDEKGFKNWAPLQ 259
QY 169 PNHI-----SAEQCASIVFWKPTGNGWNDVICETRRNSICE 204
DB 260 PDNWFHGLGGGEDCAHIT----TGGPMNDVDCQRTFRWICEM 298

RESULT 5
KUCR_RAT STANDARD; PRT; 550 AA.
AC P10716;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE C-type lectin 13 (Kupffer cell receptor).
GN CLECSF13 OR KCLR.
OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RX SEQUENCE FROM N.A., AND SEQUENCE OF 83-104.
RP MEDLINE=88227939; PubMed=2836387;
RA Hoyle G.W., Hill R.L.;
RT "Molecular cloning and sequencing of a cDNA for a carbohydrate
RT binding receptor unique to rat Kupffer cells.";
RL J. Biol. Chem. 263:7487-7492(1988).
RN [2]
RX SEQUENCE FROM N.A.
RP MEDLINE=91107689; PubMed=1846367;
RA Hoyle G.W., Hill R.L.;
RT "Structure of the gene for a carbohydrate-binding receptor unique to
RT rat Kupffer cells.";
RL J. Biol. Chem. 266:1850-1857(1991).
CC -!- FUNCTION: RECEPTOR WITH AN AFFINITY FOR GALACTOSE AND FUCOSE.
CC COULD BE INVOLVED IN ENDOCYTOSIS.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC -!- TISSUE SPECIFICITY: KUPFFER CELLS.
CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J03734; AAA41472.1; -.
DR EMBL; M55532; AAA40892.1; -.
DR PIR; A28166; A28166.
DR PIR; A38674; A38674.
DR HSSP; P20693; 1HLJ.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR00017; Syntaxin.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00503; SYN1.1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
DR Receptor; Transmembrane; Glycoprotein; Lectin; Signal-anchor;
KW Endocytosis.
FT DOMAIN 1 42 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 43 69 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT FT
FT DOMAIN 70 550
FT FT EXTRACELLULAR (POTENTIAL).
FT DISULFID 438 538 C-TYPE LECTIN (SHORT FORM).
FT DISULFID 516 528 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 93 93 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 115 115 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 132 132 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 209 209 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 550 AA; 61104 MW; 9358A6CF4C306270 CRC64;

Query Match 19.9%; Score 231.5; DB 1; Length 550;
Best Local Similarity 30.2%; Pred. No. 1.7e-13;
Matches 48; Conservative 29; Mismatches 65; Indels 17; Gaps 5;

QY 55 SELHSHSLTCTCFSEGTGVPAWGCCPASWKSFGSSCYFISSEKVKWSKSEQNCVEMGA 112
DB 386 SDLEALQKAAVAHTQGTQNTQNVQLIMQDWKFNKGYFFSRDKKSWHEAENFCVSOQA 445
QY 113 HLNVFNTEAEQNTFVQQLNESFSYFLGLSDPQGNNNWQIDKTP--YKKNVRFHILGE- 169
DB 446 HLASVTSQEQAEFLV-QITNAVVDHWIGLTDGTEGNWRWVDGTPFDYVQSRFRWKGQPD 504
QY 170 -----NHSAEQCASIVFWKPTGNGWNDVICETRRNSICE 203
```

```
RESULT 3
LECI_RAT
ID LECI_RAT STANDARD; PRT; 301 AA.
AC P08290;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Asialoglycoprotein receptor R2/3 (Hepatic lectin 2/3) (RHL-2) (ASGP-R)
DE (ASGPR).
GN ASGR2 OR ASGR-2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=87257885; PubMed=3600647;
RA McPhaul M., Berg P.;
RT "Identification and characterization of cDNA clones encoding two
RT homologous proteins that are part of the asialoglycoprotein
RT receptor.";
RL Mol. Cell. Biol. 7:1841-1847(1987).
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE=87250656; PubMed=3597443;
RA Halberg D.F., Wager R.E., Farrell D.C., Hildreth J., Quesenberry M.S.,
RA Loeb J.A., Holland E.C., Drickamer K.;
RT "Major and minor forms of the rat liver asialoglycoprotein receptor
RT are independent galactose-binding proteins. Primary structure and
RT glycosylation heterogeneity of minor receptor forms.";
RL J. Biol. Chem. 262:9828-9838(1987).
RN [3]
RN SEQUENCE FROM N.A.
RX STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=89170119; PubMed=3234178;
RA Sanford J.P., Elliott R.W., Doyle D.;
RT "Asialoglycoprotein receptor genes are linked on chromosome 11 in the
RT mouse.";
RL DNA 7:721-728(1988).
RN [4]
RN SEQUENCE OF 201-301.
RX MEDLINE=84111554; PubMed=6319386;
RA Drickamer K., Mamon J.F., Binns G., Leung J.O.;
RT "Primary structure of the rat liver asialoglycoprotein receptor.
RT Structural evidence for multiple polypeptide species.";
RL J. Biol. Chem. 259:770-778(1984).
CC -1- FUNCTION: MEDIATES THE ENDOCYTOSIS OF PLASMA GLYCOPROTEINS TO
CC WHICH THE TERMINAL SIALIC ACID RESIDUE ON THEIR COMPLEX
CC CARBOHYDRATE MOIETIES HAS BEEN REMOVED. THE RECEPTOR RECOGNIZES
CC TERMINAL GALACTOSE AND N-ACETYLGALACTOSAMINE UNITS. AFTER LIGAND
CC BINDING TO THE RECEPTOR, THE RESULTING COMPLEX IS INTERNALIZED AND
CC TRANSPORTED TO A SORTING ORGANELLE, WHERE RECEPTOR AND LIGAND ARE
CC DISASSOCIATED. THE RECEPTOR THEN RETURNS TO THE CELL MEMBRANE
CC SURFACE.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein.
CC -1- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN HEPATIC PARENCHYMAL
CC CELLS.
CC -1- MISCELLANEOUS: CALCIUM IS REQUIRED FOR LIGAND BINDING.
CC -1- MISCELLANEOUS: TWO TYPES OF RAT HEPATIC LECTIN HAVE BEEN
CC IDENTIFIED, RHL-1 AND RHL-2/3, HAVING A RELATIVE ABUNDANCE OF 4:1.
CC RHL-2 AND RHL-3 ONLY DIFFERS IN THEIR CARBOHYDRATE STRUCTURES.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; M16347; AAA42038.1; -.
EMBL; J02762; AAA41522.1; -.

```

```
DR EMBL; X07636; CAA30476.1; -.
DR PIR; A25417; LNRT2.
DR PIR; A26888; A26888.
DR PIR; A28462; A28462.
DR PIR; A31601; A31601.
DR HSSP; P06734; 1HLI.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C-TYPE_LECTIN_1; 1.
DR PROSITE; PS00441; C-TYPE_LECTIN_2; 1.
KW Lectin; glycoprotein; Receptor; Endocytosis; Transmembrane;
KW Calcium; Signal-anchor; Phosphorylation.
FT DOMAIN 1 58
FT TRANSMEM 59 79
FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT (POTENTIAL).
FT DOMAIN 80 301
FT DOMAIN 169 295
FT DISULFID 170 181
FT DISULFID 198 293
FT DISULFID 271 285
FT CARBOHYD 97 97
FT CARBOHYD 119 119
FT CARBOHYD 165 165
FT CONFLICT 153 153
FT CONFLICT 202 202
FT CONFLICT 260 260
FT CONFLICT 301 AA; 34943 MW; 3C2315E642D71279 CRC64;
SQ
Query Match 21.5%; Score 250; DB 1; Length 301;
Best Local Similarity 29.8%; Pred. No. 1.9e-15;
Matches 54; Conservative 33; Mismatches 56; Indels 38; Gaps 7;
QY 52 KRSLSELHSYHS-----SLTC-----FSEGTKVPWGCCPASKWSFGSSCYF 92
DB 129 KKQKDIKADHSTLLFHLKHPDLRLTLCQLAFFLSNGTE-----CCPVNMYVEFGSCYW 183
QY 93 ISSEKVKWSKQNCVEMGAHLVWNTAEQNFIVQQLNESFYFLGLSDPOGNNNWOWI 152
DB 184 FSRDGLTWAEADQYCOMETAHLVLSNREEQEFVVKH-RGAFHIWIGLTDKDG--SWKWV 240
QY 153 DKTPYEKNVRFWHLGEPNH-----SAPQCAASIVFWKPTGMGWNVDVICETRRNSICEMN 205
DB 241 DGTVEYRSNEKKNWAFQPDNCOGHEEGSDCAEIL-----SDGLWNDFCQQVNRWACERK 296
QY 206 K 206
DB 297 R 297
RESULT 4
MMGL_MOUSE
ID MMGL_MOUSE STANDARD; PRT; 304 AA.
AC P49300;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Macrophage asialoglycoprotein-binding protein (M-ASGP-BP) (Macrophage
DE galactose/N-acetylgalactosamine-specific lectin) (MMGL).
GN MGL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RX STRAIN=C3H/HeN;
RX MEDLINE=92268032; PubMed=1587794;
RA Sato M., Kawakami K., Osawa T., Toyoshima S.;
RT "Molecular cloning and expression of cDNA encoding a galactose/N-
RT acetylgalactosamine-specific lectin on mouse tumoricidal
RT macrophages.";
RL J. Biochem. 111:331-336(1992).

```

```

CC CC -1- SUBCELLULAR LOCATION: Type II membrane protein.
CC CC -1- ALTERNATIVE PRODUCTS: 3 isoforms; 1 (shown here), 2 and 3; are
CC CC produced by alternative splicing.
CC CC -1- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN HEPATIC PARENCHYMAL
CC CC CELLS.
CC CC -1- MISCELLANEOUS: CALCIUM IS REQUIRED FOR LIGAND BINDING.
CC CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC CC
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC
CC CC EMBL; M11025; AAB59519.1; -.
CC CC EMBL; U97197; AAB58308.1; -.
CC CC EMBL; X55283; CAA38997.1; -.
CC CC EMBL; BC017251; AAH17251.1; -.
CC CC FTR; A25179; LNHU2A.
CC CC HSP; P06734; IHLI.
CC CC Genew; HGNC:743; ASGR2.
CC CC MIM; 108361; -.
CC CC InterPro; IPR001304; Lectin_C.
CC CC Pfam; PF00059; lectin_c; 1.
CC CC SMART; SM00034; CLECT; 1.
CC CC PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
CC CC PROSITE; PS30041; C_TYPE_LECTIN_2; 1.
CC CC Lectin; glycoprotein; Receptor; Endocytosis; Transmembrane;
CC CC Calcium; Signal-anchor; Phosphorylation; Alternative splicing.
CC CC DOMAIN 1 58
CC CC TRANSMEM 59 79
CC CC SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
CC CC (POTENTIAL).
CC CC EXTRACELLULAR (POTENTIAL).
CC CC C-TYPE LECTIN (LONG FORM).
CC CC ENDOCYTOSIS SIGNAL (POTENTIAL).
CC CC BY SIMILARITY.
CC CC BY SIMILARITY.
CC CC BY SIMILARITY.
CC CC N-LINKED (GLCNAC. . .).
CC CC N-LINKED (GLCNAC. . .).
CC CC N-LINKED (GLCNAC. . .).
CC CC PHOSPHORYLATION.
CC CC MISSING (IN ISOFORM 2 AND ISOFORM 3).
CC CC MISSING (IN ISOFORM 3).
CC CC SEQUENCE 311 AA; 35191 MW; 82C78FEC8FEBA316 CRC64;
CC CC
CC CC Query Match 21.7%; Score 252.5; DB 1; Length 311;
CC CC Best Local Similarity 27.0%; Pred. No. 1.2e-15;
CC CC Matches 73; Conservative 36; Mismatches 72; Indels 89; Gaps 11;
CC CC
CC CC QY 16 LSURLNSVAGISIALLS--ACFIVSCVVTYHHTYGTGKRL-SELHSYHSSLCTPSE-- 69
CC CC Db 49 LAQRLCSWCFSLAUSFNILLVVIC-VTGSQSEGHRAQLQAEURLSKAFSNFSST 107
CC CC QY 70 -----GKTPAWG----- 77
CC CC Db 108 LTEVOALSTHGGSGVDKITSLGAKLEKQOQDLKADHALLFHLKHPVDLRFVACOWELL 167
CC CC QY 78 -----CCPASWKSFGSCYFISSEKVKWSKSEONCVEMGAHLVVFNTAEONFIVQ 129
CC CC Db 168 HSGNSQRTCCPVNVEHOGSGCYFWSHSGKAAEAERYCQLENAHLVIVNSWESQKFTVQH 227
CC CC QY 130 LNESFSYFCLGSDPQGNMNMOWIDKTPYEKNVRFHGLPENH-----SAEOCASIVFW 182
CC CC Db 228 TN-PFNTWIGLTSDG--SKWVGDGTDYRINYKNWAVTQPDNWHGHELGSQDCVEV--- 281
CC CC QY 183 KPTGWGND-----VICETRNSTICEM 204
CC CC Db 282 QPDG-RWNDQFCLQVYRWVCEKRRNATGEV 310

```

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 22:59:31 ; Search time 29 Seconds  
(without alignments)  
298.916 Million cell updates/sec

Title: US-09-766-511b-53  
Perfect score: 1165  
Sequence: 1 MMQEQQPQSTKRWLSRL.....NDVICETRRNSICENMKIYL 209

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	259	22.2	301	1	LECI_MOUSE
2	252.5	21.7	311	1	LECI_HUMAN
3	250	21.5	301	1	LECI_RAT
4	244	20.9	304	1	MMGL_MOUSE
5	231.5	19.9	550	1	KUCR_RAT
6	231	19.8	207	1	LECH_CHICK
7	226.5	19.4	283	1	LECH_RAT
8	222	19.1	306	1	MMGL_RAT
9	217	18.6	283	1	LECH_MOUSE
10	200.5	17.2	290	1	LECH_HUMAN
11	194.5	16.7	147	1	LECE_ANTCR
12	189.5	16.3	548	1	KUCR_MOUSE
13	183	15.7	883	1	PGCB_MOUSE
14	183	15.7	883	1	PGCB_RAT
15	182.5	15.7	912	1	PGCB_BOVIN
16	182	15.6	1456	1	MANR_HUMAN
17	180.5	15.5	331	1	FCE2_MOUSE
18	177.5	15.2	163	1	ANP_HEMAM
19	173.5	14.9	166	1	LITA_HUMAN
20	165	14.2	146	1	TXB_TRIFL
21	163.5	14.0	173	1	LIT2_MOUSE
22	160.5	13.8	135	1	LECG_CROAT
23	158.5	13.6	165	1	LIT1_MOUSE
24	158.5	13.6	3562	1	PGCV_CHICK
25	157.5	13.5	158	1	LECG_TRIST
26	157.5	13.5	165	1	LITH_RAT
27	154.5	13.3	166	1	LITB_HUMAN
28	153.5	13.2	3381	1	PGCV_BOVIN
29	153	13.1	175	1	ANP_OSMO
30	152	13.0	123	1	ABA4_TRIAB
31	151.5	13.0	2109	1	PGCA_CHICK
32	151.5	13.0	2738	1	PGCV_RAT
33	151.5	13.0	3358	1	PGCV_MOUSE

34	151.5	13.0	3396	1	PGCV_HUMAN
35	146	12.5	125	1	BOTB_BOTJA
36	144.5	12.4	1257	1	PGCN_RAT
37	144.5	12.4	1268	1	PGCN_MOUSE
38	143	12.3	223	1	NK12_MOUSE
39	142.5	12.2	175	1	PAP1_HUMAN
40	142	12.2	263	1	KLR4_MOUSE
41	140.5	12.1	1463	1	PA2R_BOVIN
42	140.5	12.1	2124	1	PGCA_RAT
43	140.5	12.1	2132	1	PGCA_MOUSE
44	140	12.0	321	1	FCE2_HUMAN
45	139.5	12.0	549	1	LEM2_RAT

RESULT 1

LECI_MOUSE	STANDARD;	PRT;	301 AA.
AC	P24721;		
DT	01-MAR-1992 (Rel. 21, Created)		
DT	01-MAR-1992 (Rel. 21, Last sequence update)		
DT	30-MAY-2000 (Rel. 39, Last annotation update)		
DE	Asialoglycoprotein receptor 2 (Hepatic lectin 2) (MHL-2) (ASGP-R)		
DE	(ASGPR).		
GN	ASGR2 OR ASGR-2.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6; TISSUE=Liver;		
RX	MEDLINE=91027942; PubMed=2223888;		
RA	Sanford J.P., Doyle D.;		
RT	"Mouse asialoglycoprotein receptor cDNA sequence: conservation of		
RT	receptor genes during mammalian evolution.";		
RL	Biochim. Biophys. Acta 1087:259-261(1990).		
CC	-!- FUNCTION: MEDIATES THE ENOCYTOSIS OF PLASMA GLYCOPROTEINS TO WHICH THE TERMINAL SIALIC ACID RESIDUE ON THEIR COMPLEX CARBOHYDRATE MOIETIES HAS BEEN REMOVED. THE RECEPTOR RECOGNIZES TERMINAL GALACTOSE AND N-ACETYLGLACTOSAMINE UNITS. AFTER LIGAND BINDING TO THE RECEPTOR, THE RESULTING COMPLEX IS INTERNALIZED AND TRANSPORTED TO A SORTING ORGANELLE, WHERE RECEPTOR AND LIGAND ARE DISASSOCIATED. THE RECEPTOR THEN RETURNS TO THE CELL MEMBRANE SURFACE.		
CC	-!- SUBCELLULAR LOCATION: Type II membrane protein.		
CC	-!- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN HEPATIC PARENCHYMAL CELLS.		
CC	-!- MISCELLANEOUS: CALCIUM IS REQUIRED FOR LIGAND BINDING.		
CC	-!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).		
CC	-----		
DR	EMBL; X53042; CAA37211.1; -		
DR	PIR; S13165; S13165.		
DR	HSSP; P06734; 1HLI.		
DR	MCD; MGI:88082; Asgr2		
DR	InterPro; IPR001304; Lectin_C.		
DR	PFam; PF00059; lectin_c; 1.		
DR	SMART; SM00034; CLECT; 1.		
DR	PROSITE; PS00615; C-TYPE LECTIN_1; 1.		
DR	PROSITE; PS50041; C-TYPE LECTIN_2; 1.		
KW	Lectin; Glycoprotein; Receptor; Endocytosis; Transmembrane;		
KW	Calcium; Signal-anchor; Phosphorylation.		
FT	DOMAIN 1 58 CYTOPLASMIC (POTENTIAL).		









A:Residues: 87-98 <YU2>  
C:Comment: The functioning ligand-binding unit of this receptor is thought to be at least  
C:Genetics:  
A:Gene: GDB:ASGR2; L:H2  
A:Cross-references: GDB:118755; OMIM:108361  
A:Map position: 17p13-17p11  
C:Superfamily: hepatic lectin; C-type lectin homology  
C:Keywords: alternative splicing; endocytosis; glycoprotein; lectin; liver; phosphoproteins  
F:1-311/Product: asialoglycoprotein receptor H2a #status predicted <MAT1>  
F:1-81.87-311/Product: asialoglycoprotein receptor H2b #status predicted <MAT2>  
F:1-58/Domain: intracellular #status predicted <INT>  
F:1-23.43-81.87-311/Product: asialoglycoprotein receptor H2c #status predicted <MAT3>  
F:59-78/Domain: transmembrane #status predicted <TM>  
F:79-311/Domain: extracellular #status predicted <EXT>  
F:177-300/Domain: C-type lectin homology <LCH>  
F:102.170.303/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 21.7%; Score 252.5; DB 1; Length 311;  
Best Local Similarity 27.0%; Pred. No. 6.5e+15;  
Matches 73; Conservative 36; Mismatches 72; Indels 89; Gaps 11;

QY 16 LSRLMSVAGISALLS--ACFIVSCVVYHTYTGTGKRL-SELHSYHSSUTCFS-- 69  
||| ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
Db 49 LAQRLCSMWCFSLALSFLLILLWVIC-VTGSQSEHGGAQLQAELRSLKEAFSNFSST 107  
||||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
QY 70 -----GTKVPAMG----- 77  
||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
Db 108 LTEVQALSTHGSGVDKITSLGAKLEKQQDLKADHDALLFLHKHPVDLRFVACOMELL 167  
QY 78 -----CCPASWKSGFGSCYPISSEEKWSKSEQNCVMGAHLVVFNTEABQNFTVQQ 129  
||| ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
Db 168 HNSGSRQTCCPVNVEHQSGCYWFSHGRAEAEPKYCOLENAHLVVINSWEQKFIVQH 227  
QY 130 LNESFYVELGLSDPGNNWNOWIDTKPYEKNVRFVHLGPBH-----SAEQCACTVFVW 182  
||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
Db 228 TN-PFTNTIGLTDSDG--SKMWVGDTYRHNYKNWAVTPDNWHGHGLEDGDCVFEV--- 281  
QY 183 KPTGWGMND-----VICETRRNSICEM 204  
||| ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
Db 282 QPDG-RWDDDFCLQVRYRWVCEKRNRATGEV 310

RESULT 5  
JC7595  
scavenger receptor with C-type lectin type I - human  
C:Species: Homo sapiens (man)  
C:Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 30-Jun-2001  
C:Accession: JC7595  
R:Nakamura, K.; Funakoshi, H.; Miyamoto, K.; Tokunaga, F.; Nakamura, T.  
Biochem. Biophys. Res. Commun. 280, 1028-1035, 2001  
A:Title: Molecular cloning and functional characterization of a human scavenger receptor  
A:Reference number: JC7595; MUID:21092718; PMID:11162630  
A:Contents: Placenta  
A:Accession: JC7595  
A:Molecule type: mRNA  
A:Residues: 1-742 <NAK>  
A:Cross-references: DDBJ:AB038518  
C:Comment: This receptor, a member of the scavenger receptor family, belonging to the tyrosine kinase receptor superfamily, plays an important role in recognizing infected cells and host defense. It forms a timer and plays a role in recognizing infected cells.  
C:Genetics:  
A:Gene: srcl-I  
A:Map position: 18p11.32  
C:Keywords: coiled coil; glycoprotein; transmembrane protein  
F:F1-39/Domain: cytosolic (amino-terminus) #status predicted <CVT>  
F:F16-19/Region: internalization signal YKRF  
F:F40-56/Domain: transmembrane #status predicted <TM>  
F:F57-112/Domain: extracellular #status predicted <EXT>  
F:F113-335/Domain: coiled coil #status predicted <CC>  
F:F369-384/Region: serine/threonine-rich #status predicted  
F:F443-589/Domain: collagen-like #status predicted <COL>  
F:F607-732/Domain: C-type lectin/carbohydrate recognition domain #status predicted <CRD>

Query Match 21.4%; Score 249; DB 2; Length 742:

A:Residues: 87-98 <YU2>  
C:Comment: The functioning ligand-binding unit of this receptor is thought to be at least  
C:Genetics:  
A:Gene: GDB:ASGR2; L:H2  
A:Cross-references: GDB:118755; OMIM:108361  
A:Map position: 17p13-17p11  
C:Superfamily: hepatic lectin; C-type lectin homology  
C:Keywords: alternative splicing; endocytosis; glycoprotein; lectin; liver; phosphoproteins  
F:1-311/Product: asialoglycoprotein receptor H2a #status predicted <MAT1>  
F:1-81.87-311/Product: asialoglycoprotein receptor H2b #status predicted <MAT2>  
F:1-58/Domains: intracellular #status predicted <INT>  
F:1-23.43-81.87-311/Product: asialoglycoprotein receptor H2c #status predicted <MAT3>  
F:59-78/Domains: transmembrane #status predicted <TM>  
F:79-311/Domains: extracellular #status predicted <EXT>  
F:177-300/Domains: C-type lectin homology <LCH>  
F:102.170.303/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 21.7%; Score 252.5; DB 1; Length 311;  
Best Local Similarity 27.0%; Pred. No. 6.5e+15;  
Matches 73; Conservative 36; Mismatches 72; Indels 89; Gaps 11;

QY 16 LSRLMSVAGISALLS--ACFIVSCVVHYHTYGTTGKRL-SELHSYHSSLTCSFE-- 69  
||| ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
Db 49 LAQRLCSMWCFSLALSFLLILLWVIC-VTGSQSEHGRLQAQLAKSLKEAFSNFSST 107  
||||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
QY 70 -----GTKVPAMG----- 77  
||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
Db 108 LTEVQALSTHGSGVDKITSLGAKLEKQQDLKADHDALLFLHKHPVDLRFVACOMELL 167  
QY 78 -----CCPASWKSGFGSCYPISSEEKWSKSEQNCVMGAHLVVFNTEAEQNFTVQQ 129  
||| ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
Db 168 HNSGQSORTCCPVNWEHQSGCYWFSHGRAEAEPKYCOLENAHLVVINSWEQKFIVQH 227  
QY 130 LNESFYVELGLSDPGNNWNOWIDTKPYEKNVRFWHLGPBH-----SAEQCACTIVEW 182  
||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
Db 228 TN-PFTNTIGLTDSDG--SKMWVGDTYRHNYKNWAVTPDNWHGHGLEDGEDCVVE--- 281  
QY 183 KPTGWGMND-----VICETRRNSICEM 204  
||| ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
Db 282 QPDG-RWDDDFCLQVRYRWVCERKRATGEV 310

RESULT 5  
JC7595  
scavenger receptor with C-type lectin type I - human  
C:Species: Homo sapiens (man)  
C>Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 30-Jun-2001  
C:Accession: JC7595  
R:Nakamura, K.; Funakoshi, H.; Miyamoto, K.; Tokunaga, F.; Nakamura, T.  
Biochem. Biophys. Res. Commun. 280, 1028-1035, 2001  
A>Title: Molecular cloning and functional characterization of a human scavenger receptor  
A:Reference number: JC7595; MUID:21092718; PMID:11162630  
A:Contents: Placenta  
A:Accession: JC7595  
A:Molecule type: mRNA  
A:Residues: 1-742 <NAK>  
A:Cross-references: DDBJ:AB038518  
C:Comment: This receptor, a member of the scavenger receptor family, belonging to the tyrosine kinase receptor superfamily, plays an important role in recognizing infected cells and in host defense. It forms a timer and plays a role in recognizing infected cells.  
C:Genetics:  
A:Gene: srcl-I  
A:Map position: 18p11.32  
C:Keywords: coiled coil; glycoprotein; transmembrane protein  
F:F1-39/Domains: cytosolic (amino-terminus) #status predicted <CYT>  
F:F16-19/Region: internalization signal YKRF  
F:F40-56/Domains: transmembrane #status predicted <TM>  
F:F57-112/Domains: extracellular #status predicted <EXT>  
F:F113-335/Domains: coiled coil #status predicted <CC>  
F:F369-384/Region: serine/threonine-rich #status predicted  
F:F443-589/Domains: collagen-like #status predicted <COL>  
F:F607-732/Domains: C-type lectin/carbohydrate recognition domain #status predicted <CRD>

Query Match 21.4%; Score 249; DB 2; Length 742:

A:Residues: 87-98 <YU2>  
C:Comment: The functioning ligand-binding unit of this receptor is thought to be at least  
C:Genetics:  
A:Gene: GDB:ASGR2; L:H2  
A:Cross-references: GDB:118755; OMIM:108361  
A:Map position: 17p13-17p11  
C:Superfamily: hepatic lectin; C-type lectin homology  
C:Keywords: alternative splicing; endocytosis; glycoprotein; lectin; liver; phosphoproteins  
F:1-311/Product: asialoglycoprotein receptor H2a #status predicted <MAT1>  
F:1-81/Region: intracellular #status predicted <INT>  
F:1-58/Domain: intracellular #status predicted <INT>  
F:1-23/43-81/311/Product: asialoglycoprotein receptor H2c #status predicted <MAT2>  
F:59-78/Domain: transmembrane #status predicted <TM>  
F:79-311/Domain: extracellular #status predicted <EXT>  
F:177-300/Domain: C-type lectin homology <LCH>  
F:102,170,303/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 21.7%; Score 252.5; DB 1; Length 311;  
Best Local Similarity 27.0%; Pred. No. 6,5e+15;  
Matches 73; Conservative 36; Mismatches 72; Indels 89; Gaps 11;

QY 16 LSRLMSVAGISALLS---ACFIVSCVVHYHTYGTTGKRL-SELHSYHSSITCFSE-- 69  
||| ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
Db 49 LAQRLCSMWCFSLALSFLLILLWVIC-VTGSQSEHGRLQAQLAKSLKEAFSNFSST 107  
||||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
QY 70 -----GTKVPAMG----- 77  
||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
Db 108 LTEVQALSTHGSGVDKITSLGAKLEKQQDLKADHDALLFLHKHPVDLRFVACOMELL 167  
QY 78 -----CCPASWKSGFGSCYPISSEEKWSKSEQNCVMGAHLVVFNTEAEQNFTVQQ 129  
||| ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
Db 168 HNSGQSORTCCPVNVEHQSGCYWFSHGRAEAPEKYCOLENAHLVVINSWEQKFIVQH 227  
QY 130 LNESFYVELGLSDPGNNWNOWIDTKPYEKNVRFWHLGPBH-----SAEQCACTIVEW 182  
||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
Db 228 TN-PFTNTIGLTDSDG--SKMWVGDTYRHNYKNWAVTPDNWHGHGLEDGEDCVFEV--- 281  
QY 183 KPTGCGMND-----VICETRRNSICEM 204  
||| ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||  
Db 282 QPDG-RWDDDFCLQVRYRWVCERKRATGEV 310

RESULT 5  
JC7595  
scavenger receptor with C-type lectin type I - human  
C:Species: Homo sapiens (man)  
C>Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 30-Jun-2001  
C:Accession: JC7595  
R:Nakamura, K.; Funakoshi, H.; Miyamoto, K.; Tokunaga, F.; Nakamura, T.  
Biochem. Biophys. Res. Commun. 280, 1028-1035, 2001  
A>Title: Molecular cloning and functional characterization of a human scavenger receptor  
A:Reference number: JC7595; MUID:21092718; PMID:11162630  
A:Contents: Placenta  
A:Accession: JC7595  
A:Molecule type: mRNA  
A:Residues: 1-742 <NAK>  
A:Cross-references: DDBJ:AB038518  
C:Comment: This receptor, a member of the scavenger receptor family, belonging to the tyrosine kinase receptor superfamily, plays an important role in host defense. It forms a timer and plays a role in recognizing infected cells.  
C:Genetics:  
A:Gene: srcl-I  
A:Map position: 18p11.32  
C:Keywords: coiled coil; glycoprotein; transmembrane protein  
F:F1-39/Domain: cytosolic (amino-terminus) #status predicted <CYT>  
F:F16-19/Region: internalization signal YKRF  
F:F40-56/Domain: transmembrane #status predicted <TM>  
F:F57-112/Domain: transmembrane #status predicted <TM>  
F:F113-335/Domain: extracellular #status predicted <EXT>  
F:F369-384/Region: coiled coil #status predicted <CC>  
F:F443-589/Domain: serine/threonine-rich #status predicted  
F:F607-732/Domain: collagen-like #status predicted <COL>  
F:F607-732/Domain: C-type lectin/carbohydrate recognition domain #status predicted <CRD>

Query Match 21.4%; Score 249; DB 2; Length 742:

A:Title: Mouse asialoglycoprotein receptor cDNA sequence: conservation of receptor genes  
A:Reference number: S13165; MUID:91027942; PMID:2223888  
A:Accession: S13165  
A:Molecule type: mRNA  
A:Residues: 1-301 <SAN>  
A:Cross-references: EMBL:X53042; NID:953104; PIDN:CAA37211.1; PID:g53105  
C:Superfamily: hepatic lectin; C-type lectin homology  
C:Keywords: glycoprotein; liver; Transmembrane protein  
F:170-293/Domain: C-type lectin homology <LCH>  
  
Query Match 22.2% Score 259; DB 2; Length 301;  
Best Local Similarity 33.3%; Pred. No. 1.7e-15;  
Matches 52; Conservative 30; Mismatches 50; Indels 24; Gaps 6;  
  
QY 63 SLTC-----FSEGTKVPAMGCCPASWKSFGSSCYFISSEKWSKSEQNCVEMGAHLVVF 117  
DB 154 TLFCQLAYFQSNQTE-----CCPVNVWVFGGSCYFWFSRDGLTWAEADQYCOLNAHLLVI 208  
QY 118 NTEAEQNFIVQOLNESFSYFLGLSDPOGNNQWIDTPYEKNVRFWHLGEPNH-----171  
DB 209 NSREEQDFVVKHRSQ-FHIWIGLTDKDG--SNKWVDGTDYRSNYRNWAFQPDNMQGHEQ 265  
QY 172 -SAEQCASIVFWKPTGCGWGNVDVICETRRNSICEMNK 206  
DB 266 GGEDCAEIL-----SDGHWNDNFCQOVNRWCEKRR 297  
  
RESULT 3  
LNRT2  
hepatic lectin 2 - rat  
N:Alternate names: asialoglycoprotein receptor RHL-2/3 (ASGP-R2/3)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 31-Mar-1988 #sequence\_revision 09-Apr-1998 #text\_change 22-Jun-1999  
C:Accession: B28462; A28462; A31601; A26888; A25417  
R:Halberd, D.F.; Weger, R.E.; Farrell, D.C.; Hildreth IV, J.; Quesenberry, M.S.; Loeb, J.  
J. Biol. Chem. 262, 9828-9838, 1987  
A:Title: Major and minor forms of the rat liver asialoglycoprotein receptor are independent  
A:Reference number: A28462; MUID:87250656; PMID:3597443  
A:Accession: B28462  
A:Molecule type: mRNA  
A:Residues: 1-301 <HAL>  
A:Cross-references: GB:J02762; NID:g205162; PIDN:AAA41522.1; PID:g205163  
A:Accession: A28462  
A:Molecule type: protein  
A:Residues: 88-96,'X',98-118,'X',120-129-158-177-182,'X',184,'X',186-189;192-290,'C',292-294  
R:Sanford, J.P.; Elliott, R.W.; Doyle, D.  
DNA 7, 721-728, 1988  
A:Title: Asialoglycoprotein receptor genes are linked on chromosome 11 in the mouse.  
A:Reference number: A31601; MUID:89170119; PMID:3234178  
A:Accession: A31601  
A:Molecule type: mRNA  
A:Residues: 1-301 <SAN>  
A:Cross-references: GB:X07636; NID:g57066; PIDN:CAA30476.1; PID:g57067  
R:McPhaul, M.; Berg, P.  
Mol. Cell. Biol. 7, 1841-1847, 1987  
A:Title: Identification and characterization of cDNA clones encoding two homologous proteins  
A:Reference number: A26888; MUID:87257885; PMID:3600647  
A:Accession: A26888  
A:Molecule type: mRNA  
A:Residues: 1-152,'A',154-201,'I',203-259,'C',261-301 <MCP>  
A:Cross-references: GB:M16347; NID:g206648; PIDN:AAA42038.1; PID:g206649  
A:Note: the authors translated the codon GCA for residue 153 as Arg and ATT for residue 170  
R:Drickamer, K.; Mamon, J.F.; Binns, G.; Leung, J.O.  
J. Biol. Chem. 259, 770-778, 1984  
A:Title: Primary structure of the rat liver asialoglycoprotein receptor: structural evidence  
A:Reference number: A25417; MUID:84111554; PMID:6319386  
A:Accession: A25417  
A:Molecule type: protein  
A:Residues: 201-259,'C',261-281,'ND',284-301 <DRI>  
C:Comment: Calcium is required for ligand binding.  
C:Superfamily: hepatic lectin; C-type lectin homology  
C:Keywords: endocytosis; glycoprotein; lectin; liver; receptor; transmembrane protein  
F:2-60/Domain: intracellular #status predicted <INT>

F:61-77/Domain: transmembrane #status predicted <TRN>  
F:78-301/Domain: extracellular #status predicted <EXT>  
F:170-293/Domain: C-type lectin homology <LCH>  
F:97,119,165/Binding site: carbohydrate (Asn) (covalent) #status predicted  
  
Query Match 21.8% Score 254; DB 1; Length 301;  
Best Local Similarity 29.8%; Pred. No. 4.6e-15;  
Matches 54; Conservative 33; Mismatches 56; Indels 38; Gaps 7;  
  
QY 52 KRUSELSHYHS-----SITC-----FSEGTKVPAMGCCPASWKSFGSSCYF 92  
DB 129 KKQDKIKADHSTLLFLHKHFPDLRLTLCOLAFSLNGTE-----CCPVNVWVFGGSCYW 183  
QY 93 ISSEKWSKSEQNCVEMGAHLVVFNTAEQNFIVQOLNESFSYFLGLSDPOGNNQWIDTPYEKNVRFWHLGEPNH-----152  
DB 184 FSRDGLTWAEADQYCOLNAHLVINSREDFEYVVKH-RGAFHIIWIGLTDKDG--SNKWV 240  
QY 153 DKTYEKNVRFWHLGEPNH-----SAEQCASIVFWKPTGCGWGNVDVICETRRNSICEMN 205  
DB 241 DGYEYRSNFKNWAFTQPDNMQGHEEGSEDCAEIL-----SDGLWNDNFCQOVNRWACERK 296  
QY 206 K 206  
DB 297 R 297  
  
RESULT 4  
LNHU2A  
asialoglycoprotein receptor H2a - human  
N:Alternate names: hepatic lectin H2a  
C:Species: Homo sapiens (man)  
C:Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 10-Dec-1999  
C:Accession: A25179; A39100; B39100; I37995; A49466; B49466; S14525  
R:Spies, M.; Lodish, H.F.  
Proc. Natl. Acad. Sci. U.S.A. 82, 6465-6469, 1985  
A:Title: Sequence of a second human asialoglycoprotein receptor: conservation of two  
A:Reference number: A25179; MUID:86016723; PMID:3863106  
A:Accession: A25179  
A:Molecule type: mRNA  
A:Residues: 1-311 <SPI>  
A:Cross-references: GB:M11025; NID:gi79080; PIDN:AAB59519.1; PID:gi79081  
R:Lederkremer, G.Z.; Lodish, H.F.  
J. Biol. Chem. 266, 1237-1244, 1991  
A:Title: An alternatively spliced minixon alters the subcellular fate of the human a  
A:Reference number: A39100; MUID:91093236; PMID:1985943  
A:Accession: A39100  
A:Molecule type: DNA; mRNA  
A:Residues: 69-99 <LED>  
A:Cross-references: GB:M38420; NID:gi184395  
A:Accession: B39100  
A:Molecule type: DNA; mRNA  
A:Residues: 69-81,87-99 <LE2>  
A:Cross-references: GB:M38420; NID:gi184395  
R:Paletta, E.; Stockert, R.J.; Racevskis, J.  
Hepatology 15, 395-402, 1992  
A:Title: Differences in the abundance of variably spliced transcripts for the second  
A:Reference number: I37995; MUID:92184202; PMID:1371982  
A:Accession: I37995  
A:Status: preliminary;  
A:Molecule type: mRNA  
A:Residues: 1-23 43-81,87-311 <PAT>  
A:Cross-references: EMBL:X55283; NID:g34354; PIDN:CAA38997.1; PID:g34355  
R:Yuk, M.H.; Lodish, H.F.  
J. Cell Biol. 123, 1735-1749, 1993  
A:Title: Two pathways for the degradation of the H2 subunit of the asialoglycoprotein  
A:Reference number: A49466; MUID:94103329; PMID:8276894  
A:Accession: A49466  
A:Molecule type: protein  
A:Residues: 78-98 <YUK>  
A:Accession: B49466  
A:Molecule type: protein

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 22:56:31 ; Search time 20 seconds  
(without alignments)  
1004.605 Million cell updates/sec

Title: US-09-766-511b-53  
Perfect score: 1165  
Sequence: 1 MMQEQQPQSTKRGWLSRL.....NDVICETRRNSICEMNKIYL 209

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PTR\_73:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	458	39.3	237	2 JC7608	Type II lectin-like
2	259	22.2	301	2 SL13165	asialoglycoprotein
3	254	21.8	301	1 LNRT2	hepatic lectin 2 -
4	252.5	21.7	311	1 LNH02A	asialoglycoprotein
5	249	21.4	742	2 JC7595	scavenger receptor
6	244	20.9	304	2 JX0209	lectin, galactose/
7	235	20.2	404	2 A46274	HIV gp120-binding
8	231.5	19.9	550	2 A28166	Kupffer cell recep
9	231	19.8	207	1 LNCHL	hepatic lectin - c
10	226.5	19.4	284	1 LNRTL	hepatic lectin - r
11	222	19.1	306	2 A42230	lectin M-AsGP-Bp p
12	221	19.0	284	2 S29855	asialoglycoprotein
13	200.5	17.2	291	1 LNHU1	hepatic lectin H1
14	194.5	16.7	147	2 A26697	echinoidin - sea u
15	189.5	16.3	1479	2 T42710	mannose receptor,
16	183	15.7	883	2 S49126	brevican precursor
17	183	15.7	883	2 S57653	brevican precursor
18	182.5	15.7	912	2 A54423	brevican precursor
19	182	15.6	1456	1 A36563	mannose receptor p
20	181.5	15.6	309	1 S34198	TgE Fc receptor II
21	180.5	15.5	330	2 T46256	brevican - human (
22	180.5	15.5	331	1 LNM5ER	IgE Fc receptor, I
23	177.5	15.2	163	1 A34313	antifreeze protein
24	176.5	15.2	1455	1 A48925	mannose receptor p
25	173.5	14.9	166	1 RGHU1A	regenerating islet
26	169.5	14.5	156	2 T28141	C type lectin, B 1
27	169	14.5	1723	2 S58880	receptor DEC-205 -
28	167.5	14.4	188	2 I59421	mast cell function
29	165	14.2	146	2 JC4691	coagulation factor

30 164.5 14.1 166 2 A45751 pancreatic stone p  
31 163.5 14.0 173 2 B47148 reg II, regenerati  
32 160.5 13.8 135 2 A38609 lectin, galactose-  
33 159.5 13.7 257 2 I50146 gene 17.5 protein  
34 158.5 13.6 165 2 A47148 reg I, regeneratin  
35 158.5 13.6 3562 2 A47171 chondroitin sulfat  
36 157.5 13.5 165 2 A28351 pancreatic stone p  
37 154.5 13.3 166 1 RGHU1B regenerating islet  
38 153.5 13.2 1643 2 T14274 versican precursor  
39 153.5 13.2 3381 2 T42389 versican precursor  
40 153 13.1 175 1 JH0626 antifreeze protein  
41 151.5 13.0 2109 1 I50421 aggrecan precursor  
42 151.5 13.0 2397 1 A55535 versican precursor  
43 151.5 13.0 2409 1 A60979 versican precursor  
44 148 12.7 155 2 S78774 perlucin - Haloti  
45 146 12.5 125 2 B47267 botroctetin beta ch

RESULT 1  
JC7608  
type II lectin-like immunoreceptor - human  
C:Species: Homo sapiens (man)  
C:Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 30-Jun-2001  
C:Accession: JC7608  
R:Huang, X.; Yuan, Z.; Chen, G.; Zhang, M.; Zhang, W.; Yu, Y.; Cao, X.  
Biochem. Biophys. Res. Commun. 281, 131-140, 2001  
A:Title: Cloning and characterization of a novel ITIM containing lectin-like immunore  
A:Reference number: JC7608; MUID:21092797; PMID:11178971  
A:Contents: Dendritic cells  
A:Accession: JC7608  
A:Molecule type: mRNA  
A:Residues: 1-237 <HUA>  
A:Cross-references: GB:AF067800  
C:Comment: This receptor, highly homologous to the group of macrophage/hepatic lectin  
cell, especially in migrating, antigen capturing and processing.  
C:Genetics:  
A:Gene: liir  
A:Map position: 12p13  
F:45-69/Domain: transmembrane #status predicted <TMM>

Query Match 39.3%; Score 458; DB 2; Length 237;  
Best Local Similarity 44.4%; Pred. No. 3e-33;  
Matches 83; Conservative 34; Mismatches 66; Indels 4; Gaps 3;

QY 26 ISTALLSACFIYSCVYHYFTYGTGKRLSELHSHSLTCTFSEGTKV--PAWGCCPASW 83  
DB 52 IFFLLLAISFFIAFVFFQ-KYSQLLEKKTRELVTITLECVKKNMPVEETAMSCCPKNW 110  
QY 84 KPSGSSCYFTSSPEKYVSKSEQNCVEMGAHVVFNTAEQNFIVQOLNESFSYFLGLSDP 143  
DB 111 KPSSSCYFTISTDSASWQSEKDCARWEAHLVINTQEEQDFIFQLQESAYFVGLSDP 170  
QY 144 QGNNNQWQDKTPEYKKNRFRVHLGEPNHSAEQCASTVFVK-PTGNGWNDVICETRRNSTC 202  
DB 171 EGQRHWQWVDQTPYNSSIFWHPREPDPNERCVLVNFRKSPKRWGNDVNCILGPQRSVC 230  
QY 203 EMNKIYL 209  
DB 231 EMNKIHL 237

RESULT 2  
SL13165  
asialoglycoprotein receptor - mouse  
N:Alternate names: hepatic lectin  
C:Species: Mus musculus (house mouse)  
C:Date: 21-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 20-Aug-1999  
C:Accession: SL13165  
R:Sanford, J.P.; Doyle, D.  
Biochim. Biophys. Acta 1087, 259-261, 1990

ALIGNMENTS

Db 182 MRDSSNPRONWNDVTCLNYFRICEM 207

Search completed: February 11, 2003, 23:02:14  
Job time : 11 secs

```
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 377
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-904-011-377

Query Match          34.2%; Score 398.5; DB 9; Length 219;
Best Local Similarity 40.8%; Pred. No. 1.1e-32;
Matches 84; Conservative 34; Mismatches 81; Indels 7; Gaps 5;

QY 3 QEQQPQSTKRGWLS--LRLMSVAGISIALLSACFIVSCVVTYHYFTYGETGKRLSELHSY 60
Db 5 KSETQCTE-RGCFSSQMFMTVAGIPILFLSACFITRCVVTFRF-FQTCDEKKFQLPEN 62
QY 61 HSSLTCFSEGTKVPAGWCCPASWKSFGSSCYFISSEKVKWSKSEQNCVEMGAHLVVFENTE 120
Db 63 FTLESCYNYGSG--SVKNCCPLNWEYFQSSCYFFSTDITISWALSCLKNCSSAMGAHLVINSQ 121
QY 121 AEQNFIVQQLNESFSYFLGLSDPQGNNNQWIDKTPYEKNVRFVHILGEPNHS--EQCAS 178
Db 122 EEQEFLSYKKPKMREFFIGLSQVVGOWVDGTPLTKSLSEFWDVGEPPNIIATLEDCA 181
QY 179 IVFWKPTGWNVDVICETRNSICEM 204
Db 182 MRDSSNPRQWNDVTCFLNYFRICEM 207

RESULT 14
US-10-174-590-24
; Sequence 24, Application US/10174590
; Publication No. US20030008352A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C42
; CURRENT APPLICATION NUMBER: US/10/174,590
; CURRENT FILING DATE: 2002-06-18
; Prior application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 377
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-590-24

Query Match          34.2%; Score 398.5; DB 9; Length 219;
Best Local Similarity 40.8%; Pred. No. 1.1e-32;
Matches 84; Conservative 34; Mismatches 81; Indels 7; Gaps 5;

QY 3 QEQQPQSTKRGWLS--LRLMSVAGISIALLSACFIVSCVVTYHYFTYGETGKRLSELHSY 60
Db 5 KSETQCTE-RGCFSSQMFMTVAGIPILFLSACFITRCVVTFRF-FQTCDEKKFQLPEN 62
QY 61 HSSLTCFSEGTKVPAGWCCPASWKSFGSSCYFISSEKVKWSKSEQNCVEMGAHLVVFENTE 120
Db 63 FTLESCYNYGSG--SVKNCCPLNWEYFQSSCYFFSTDITISWALSCLKNCSSAMGAHLVINSQ 121
QY 121 AEQNFIVQQLNESFSYFLGLSDPQGNNNQWIDKTPYEKNVRFVHILGEPNHS--EQCAS 178
Db 122 EEQEFLSYKKPKMREFFIGLSQVVGOWVDGTPLTKSLSEFWDVGEPPNIIATLEDCA 181
QY 179 IVFWKPTGWNVDVICETRNSICEM 204
Db 182 MRDSSNPRQWNDVTCFLNYFRICEM 207

RESULT 15
US-10-176-758-24
; Sequence 24, Application US/10176758
; Publication No. US20030008353A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C104
; CURRENT APPLICATION NUMBER: US/10/176,758
; CURRENT FILING DATE: 2002-06-21
; Prior application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 24
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-758-24

Query Match          34.2%; Score 398.5; DB 9; Length 219;
Best Local Similarity 40.8%; Pred. No. 1.1e-32;
Matches 84; Conservative 34; Mismatches 81; Indels 7; Gaps 5;

QY 3 QEQQPQSTKRGWLS--LRLMSVAGISIALLSACFIVSCVVTYHYFTYGETGKRLSELHSY 60
Db 5 KSETQCTE-RGCFSSQMFMTVAGIPILFLSACFITRCVVTFRF-FQTCDEKKFQLPEN 62
QY 61 HSSLTCFSEGTKVPAGWCCPASWKSFGSSCYFISSEKVKWSKSEQNCVEMGAHLVVFENTE 120
Db 63 FTLESCYNYGSG--SVKNCCPLNWEYFQSSCYFFSTDITISWALSCLKNCSSAMGAHLVINSQ 121
QY 121 AEQNFIVQQLNESFSYFLGLSDPQGNNNQWIDKTPYEKNVRFVHILGEPNHS--EQCAS 178
Db 122 EEQEFLSYKKPKMREFFIGLSQVVGOWVDGTPLTKSLSEFWDVGEPPNIIATLEDCA 181
QY 179 IVFWKPTGWNVDVICETRNSICEM 204
; Prior application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 24
```



Db 63 FTSLSCYNYGSG-SVKNCCPLNWYFQSSCYFFSTDTTISWALSILKNCSAMGAHLVINSQ 121  
QY 121 AEONFTVOOLNESFYFLGLSDPOGNNWQIDKTPYEKNVRFWHLGEPNHS--EOCAS 178  
Db 122 EGEFELSYKKPKMREFFGLSDQVVEGQWVDGTPLTLSFWDVGEPPNIIATLEDCAT 181  
QY 179 IVFWKPTGWNVDVICETRNSICEM 204  
Db 182 MRDSSNPQWNDVTCFLNYFRICEM 207

## RESULT 12

US-09-907-841-377

; Sequence 377, Application US/09907841  
; Publication No. US20020198366A1  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc.  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, A.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Hillan, Kenneth, J.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Mather, Jennie P.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William, I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: 10466-14  
; CURRENT APPLICATION NUMBER: US/09/907,841  
; PRIOR FILING DATE: 2001-11-20  
; PRIOR APPLICATION NUMBER: PCT/US00/04414  
; PRIOR FILING DATE: 2000-02-22  
; PRIOR FILING DATE: 1999-07-07  
; PRIOR FILING DATE: 1999-07-26  
; PRIOR FILING DATE: 1999-07-26  
; PRIOR FILING DATE: 1999-07-28  
; PRIOR FILING DATE: 1999-09-08  
; PRIOR FILING DATE: 1999-09-13  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR FILING DATE: 1999-10-29  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 423  
; SEQ ID NO 377  
; LENGTH: 219  
; TYPE: PRT  
; ORGANISM: Homo Sapien

US-09-907-841-377

Query Match 34.2%; Score 398.5; DB 9; Length 219;  
Best Local Similarity 40.8%; Pred. No. 1.le-32;  
Matches 84; Conservative 34; Mismatches 81; Indels 7; Gaps 5;  
QY 3 QOQOQSQSTKRCWL--LRLSVAGISALISACFIIVSCVVTYHTYGTGKRLSELHSY 60  
Db 5 KSETQCTE-RGCFSSQMFVAGIPILFSLACFITRCVVTFRI-FQTCDEKFKFQLEN 62  
QY 61 HSLTCTFSEGTKVPWAGCCPASWGSFSGSCYFISSEKVKWSQNCVEMGAHLVVFENTE 120  
Db 63 FTSLSCYNYGSG-SVKNCCPLNWYFQSSCYFFSTDTTISWALSILKNCSAMGAHLVINSQ 121  
QY 121 AEONFTVOOLNESFYFLGLSDPOGNNWQIDKTPYEKNVRFWHLGEPNHS--EOCAS 178  
Db 122 EGEFELSYKKPKMREFFGLSDQVVEGQWVDGTPLTLSFWDVGEPPNIIATLEDCAT 181  
QY 179 IVFWKPTGWNVDVICETRNSICEM 204  
Db 182 MRDSSNPQWNDVTCFLNYFRICEM 207

## RESULT 13

US-09-904-011-377

; Sequence 377, Application US/09904011  
; Publication No. US20030003530A1  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc.  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, A.  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth, J.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Mather, Jennie P.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William, I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: 10466-14  
; CURRENT APPLICATION NUMBER: US/09/904,011  
; CURRENT FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 09/665,350  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: PCT/US00/04414  
; PRIOR FILING DATE: 2000-02-22  
; PRIOR APPLICATION NUMBER: US 60/143,048  
; PRIOR FILING DATE: 1999-07-07  
; PRIOR APPLICATION NUMBER: US 60/145,698  
; PRIOR FILING DATE: 1999-07-26  
; PRIOR APPLICATION NUMBER: US 60/146,222  
; PRIOR FILING DATE: 1999-07-28  
; PRIOR APPLICATION NUMBER: PCT/US99/20594  
; PRIOR FILING DATE: 1999-09-08  
; PRIOR APPLICATION NUMBER: PCT/US99/20944  
; PRIOR FILING DATE: 1999-09-13  
; PRIOR APPLICATION NUMBER: PCT/US99/21090

```
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 377
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-09-902-853-377

Query Match          34.2%; Score 398.5; DB 9; Length 219;
Best Local Similarity 40.8%; Pred. No. 1.1e-32;
Matches 84; Conservative 34; Mismatches 81; Indels 7; Gaps 5;

QY 3 QEQQPQSTQKRGWLS--LRLMSVAGISIALLSACFIVSCVVTYHFTYGTGKRLSELHSY 60
Db 5 KSETQCTE-RGCFSSOMFLWTAGIPILFSLACFTITRCVVTFRI-FQTCDEKKFQLPEN 62
QY 61 HSSLTFCSEGTQKVPANGCCPASWKSFGSSCYFISSEKVKWSKSEQNCQVEMGAHLVVFNTE 120
Db 63 FTLSVCYNYGSG-SVKNCCPLNWEYFQSSCYFTFTDTISWALS LKNCSSAMGAHLVYVINSQ 121
QY 121 AEQNFIVQOLNESFSYFLGLSDPQGNNNQWIDKTPYKVRVHGLGEPNHSAA--EQCAS 178
Db 122 EEQEFLSYKKPKMREFFIGLSDQVQGVQWQWVDGTPLTSLSFWDVGFENNIATLEDCAT 181
QY 179 IVFWKPTGWNVDVICETRNSTCEM 204
Db 182 MRDSSNPRQNWDVTCLNFRICEM 207

RESULT 11
US-09-907-824-377
; Publication 377, Application US/09907824
; Publication No. US20020197671A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Wei-Oliang
; APPLICANT: Gao, Wei-Oliang
```

```

; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 377
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-905-291A-377
```

```

Query Match          34.2%; Score 398.5; DB 9; Length 219;
Best Local Similarity 40.8%; Pred. No. 1.le-32;
Matches 84; Conservative 34; Mismatches 81; Indels 7; Gaps 5;

QY 3 QEQQPQSTKRGWLS--LRLMSVAGISIALLSACFIVSCVVTYHFTYGETCKRLSELHSY 60
   : : | | | | : : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 5 KSSETQCTE-RGCFSSQMFMTWAGIPILFLSACFITRCVVTFR1-FQTCDEKKFQLPEN 62
   : | | : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 61 HSSLTCESEGTKVPAMGCCPASKWKSFGSSCYFISSEKVKWSEONCVEMGAHLVVFNTE 120
   : | | : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 63 FTLSYNYGSG-SVKNCCPLNWEYFQSSCYFFSTDTISWLSLKNCSAMGAHLVWINSQ 121
   || | : : : : | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 121 AEONFIVQOLNESFSYFLGLSDPQGNMNMOWIDKTPYEKNVRFWHLGEPNHS1A--EQCAS 178
   || | : : : : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 122 EQEFLSYKKPKMREFFIGLSQDVVEGQWQVGDGTPLTKSLSFWDVGEPPN1ATLEDCAT 181
   || | : : : : | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 179 IVFWKPTGWNNDVICETRRNSICEM 204
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 182 MRDSSNPRQWNDVTCFLNYFRICEM 207
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```

RESULT 9
US-09-965-529-3
; Sequence 3, Application US/09965529
; Publication No. US20020182671A1
; GENERAL INFORMATION:
; APPLICANT: LAL, Preeti
; APPLICANT: YUE, Henry
; APPLICANT: TANG, Y. Tom
; APPLICANT: BANDMAN, Olga
; APPLICANT: BURFORD, Neil
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0731 USA
; CURRENT APPLICATION NUMBER: US/09/965,529
```

```

; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/149,641; 60/164,203; PCT/US00/22315
; PRIOR FILING DATE: 1999-08-17; 1999-11-09; 2000-08-14
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020182671A1 1521513CD1
US-09-965-529-3

Query Match          34.2%; Score 398.5; DB 9; Length 219;
Best Local Similarity 40.8%; Pred. No. 1.le-32;
Matches 84; Conservative 34; Mismatches 81; Indels 7; Gaps 5;

QY 3 QEQQPQSTKRGWLS--LRLMSVAGISIALLSACFIVSCVVTYHFTYGETCKRLSELHSY 60
   : : | | | | : : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 5 KSSETQCTE-RGCFSSQMFMTWAGIPILFLSACFITRCVVTFR1-FQTCDEKKFQLPEN 62
   : | | : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 61 HSSLTCESEGTKVPAMGCCPASKWKSFGSSCYFISSEKVKWSEONCVEMGAHLVVFNTE 120
   : | | : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 63 FTLSYNYGSG-SVKNCCPLNWEYFQSSCYFFSTDTISWLSLKNCSAMGAHLVWINSQ 121
   || | : : : : | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 121 AEONFIVQOLNESFSYFLGLSDPQGNMNMOWIDKTPYEKNVRFWHLGEPNHS1A--EQCAS 178
   || | : : : : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 122 EQEFLSYKKPKMREFFIGLSQDVVEGQWQVGDGTPLTKSLSFWDVGEPPN1ATLEDCAT 181
   || | : : : : | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 179 IVFWKPTGWNNDVICETRRNSICEM 204
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 182 MRDSSNPRQWNDVTCFLNYFRICEM 207
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```

RESULT 10
US-09-902-853-377
; Sequence 377, Application US/09902853
; Publication No. US20020192659A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Baton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/902,853
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US/09/665,350
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US 60/143,048
```

Query Match	39.3%;	Score 458;	DB 10;	Length 243;
Best Local Similarity	44.4%;	Pred. No. 1.3e-38;		
Matches	83;	Conservative	34;	Mismatches 66; Indels 4; Gaps 3;

  

QY	26	ISTALLSACFIVSCVVYHTFYGTGKRLSELHSHYSSLTLCFSEGTKV--PANGCCCPASW	83
Db	58	IFFLLLAISFFIAFVIFFO-KYSQLEKTKTKELVHTTLECVKKNMPVEETAWSOCPKNW	116
QY	84	KSFSSCYFISSEKVKSKSEONCVEMGAHLVVFNTAEONFTVOOLNESFSYFLGLSDP	143
Db	117	KSFSSNCYFISTESASQDSEKDCARMAHLVINTQEOQDFIQNLQESAYFVGLSDP	176
QY	144	QGNNNOWIDKTPYEKNVRFWHLGEPNHSABQCASIVFWK-PTGWGNDVICETRNNSIC	202
Db	177	EGQRHWQVDQTPYNESSTFWHPREPSDPNERCVVLNFRKSPKRWGNDVNCILGPORSVC	236
QY	203	EMMKIYL 209	
Db	237	EMMKIHL 243	

  

RESULT 6

US-09-764-870-303

; Sequence 303, Application US/09764870

; Patent No. US20020042386A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PTZ14

; CURRENT APPLICATION NUMBER: US/09/764,870

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - consult PALM or file wrapper

; NUMBER OF SEQ ID NOS: 646

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 303

; LENGTH: 246

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-764-870-303

  

Query Match	39.3%;	Score 458;	DB 10;	Length 246;
Best Local Similarity	44.4%;	Pred. No. 1.3e-38;		
Matches	83;	Conservative	34;	Mismatches 66; Indels 4; Gaps 3;

  

QY	26	ISTALLSACFIVSCVVYHTFYGTGKRLSELHSHYSSLTLCFSEGTKV--PANGCCCPASW	83
Db	61	IFFLLLAISFFIAFVIFFO-KYSQLEKTKTKELVHTTLECVKKNMPVEETAWSOCPKNW	119
QY	84	KSFSSCYFISSEKVKSKSEONCVEMGAHLVVFNTAEONFTVOOLNESFSYFLGLSDP	143
Db	120	KSFSSNCYFISTESASQDSEKDCARMAHLVINTQEOQDFIQNLQESAYFVGLSDP	179
QY	144	QGNNNOWIDKTPYEKNVRFWHLGEPNHSABQCASIVFWK-PTGWGNDVICETRNNSIC	202
Db	180	EGQRHWQVDQTPYNESSTFWHPREPSDPNERCVVLNFRKSPKRWGNDVNCILGPORSVC	239
QY	203	EMMKIYL 209	
Db	240	EMMKIHL 246	

  

RESULT 7

US-09-862-802-8

; Sequence 8, Application US/09862802

; Patent No. US20020165346A1

; GENERAL INFORMATION:

; APPLICANT: Schering-Plough Corporation

; TITLE OF INVENTION: MAMMALIAN MEMBRANE PROTEIN GENES; RELATED REAGENTS

; FILE REFERENCE: SF0699B

; CURRENT APPLICATION NUMBER: US/09/862,802

; CURRENT FILING DATE: 2002-01-10

; PRIOR APPLICATION NUMBER: US 09/111,470